SYMANTEC ENTERPRISE SECURITY

Confidence in a connected world.

Symantec Internet Security Threat Report

Trends for January–June 07

Volume XII, Published September 2007

Dean Turner
Executive Editor
Symantec Security Response

Stephen Entwisle
Senior Editor
Symantec Security Response

Eric Johnson
Editor
Symantec Security Response

Marc Fossi
Analyst
Symantec Security Response

Joseph Blackbird
Analyst
Symantec Security Response

David McKinney
Analyst
Symantec Security Response

Ronald Bowes
Analyst
Symantec Security Response

Nicholas Sullivan
Analyst
Symantec Security Response

Candid Wueest
Analyst
Symantec Security Response

Ollie Whitehouse
Security Architect—Advanced Threat Research
Symantec Security Response

Zulfikar Ramzan
Analyst—Advanced Threat Research
Symantec Security Response

Jim Hoagland
Principal Software Engineer
Symantec Security Response

Chris Wee
Manager, Development
Symantec Security Response

Contributors

David Cowings
Sr. Manager of Operations
Symantec Business Intelligence

Dylan Morss
Manager
Symantec Business Intelligence

Shravan Shashikant
Principal Business Intelligence Analyst
Symantec Business Intelligence

Volume XII, Published September 2007

Symantec Internet Security Threat Report

Contents

Symantec *Internet Security Threat Report* Executive Summary

The Symantec *Internet Security Threat Report* provides a six-month update of Internet threat activity. It includes analysis of network-based attacks, a review of known vulnerabilities, and highlights of malicious code. It will also assess trends in phishing and spam activity. This summary of the *Internet Security Threat Report* will alert readers to current trends and impending threats. It will also offer recommendations for protection against and mitigation of these concerns. This volume covers the six-month period from January 1 to June 30, 2007.

Symantec has established some of the most comprehensive sources of Internet threat data in the world. The Symantec™ Global Intelligence Network tracks attack activity across the entire Internet. It consists of over 40,000 sensors monitoring network activity in over 180 countries. As well, Symantec gathers malicious code reports from over 120 million client, server, and gateway systems that have deployed Symantec's antivirus products.

Symantec operates one of the most popular forums for the disclosure and discussion of vulnerabilities on the Internet, the BugTraq™ mailing list, which has approximately 50,000 direct subscribers who contribute, receive, and discuss vulnerability research on a daily basis.[1] Symantec also maintains one of the world's most comprehensive vulnerability databases, currently consisting of over 22,000 vulnerabilities (spanning more than a decade) affecting more than 50,000 technologies from over 8,000 vendors. The following discussion of vulnerability trends is based on a thorough analysis of that data.

Finally, the Symantec Probe Network, a system of over two million decoy accounts, attracts email messages from 20 different countries around the world, allowing Symantec to gauge global spam and phishing activity. These resources give Symantec analysts unparalleled sources of data with which to identify emerging trends in attacks and malicious code activity. Symantec also gathers phishing information through the Symantec Phish Report Network, an extensive antifraud community of enterprises and consumers. Members of the network contribute and receive fraudulent Web site addresses for alerting and filtering across a broad range of solutions.

The Symantec *Internet Security Threat Report* is grounded principally on the expert analysis of data provided by all of these sources. Based on Symantec's expertise and experience, this analysis yields a highly informed commentary on current Internet threat activity. By publishing the Symantec *Internet Security Threat Report*, Symantec hopes to provide enterprises and consumers with the information they need to help effectively secure their systems now and in the future.

Executive Summary Highlights

The following section will offer a brief summary of the security trends that Symantec observed during this period based on data provided by the sources listed above. This summary includes all of the metrics that are included in the main report. Following this overview, the Executive Summary of the *Internet Security Threat Report* will discuss selected metrics in greater depth.

[1] The BugTraq mailing list is hosted by SecurityFocus (http://www.securityfocus.com). Archives are available at http://www.securityfocus.com/archive/1

Attack Trends Highlights

- The United States was the country targeted by the most denial of service (DoS) attacks, accounting for 61 percent of the worldwide total in the first half of 2007.
- The United States was the top country of attack origin in the first six months of 2007, accounting for 25 percent of the worldwide attack activity.
- During this period, the United States accounted for 30 percent of all malicious activity during the period, more than any other country.
- Israel was the country with the most malicious activity per Internet user in the first six months of 2007, followed by Canada and the United States.
- Four percent of all malicious activity detected during the first six months of 2007 originated from IP space registered to Fortune 100 companies.
- The education sector accounted for 30 percent of data breaches that could lead to identity theft during this period, more than any other sector.
- Theft or loss of computer or other data-storage medium made up 46 percent of all data breaches that could lead to identity theft during this period.
- The United States was the top country for underground economy servers, accounting for 64 percent of the total known to Symantec.
- Credit cards were the most common commodity advertised on underground economy servers known to Symantec, accounting for 22 percent of all items.
- Eighty-five percent of credit cards advertised for sale on underground economy servers known to Symantec were issued by banks in the United States.
- Symantec observed an average of 52,771 active bot-infected computers per day in the first half of 2007, a 17 percent decrease from the previous period.
- China had 29 percent of the world's bot-infected computers, more than any other country.
- The United States had the highest number of bot command-and-control servers, accounting for 43 percent of the worldwide total.
- Beijing was the city with the most bot-infected computers, accounting for seven percent of the worldwide total.
- The average lifespan of a bot-infected computer during the first six months of 2007 was four days, up from three days in the second half of 2006.
- Home users were the most highly targeted sector, accounting for 95 percent of all targeted attacks.

Vulnerability Trends Highlights

- Symantec documented 2,461 vulnerabilities in the first half of 2007, three percent less than the second half of 2006.
- Symantec classified nine percent of all vulnerabilities disclosed during this period as high severity, 51 percent were medium severity, and 40 percent were low. In the second half of 2006, four percent of newly disclosed vulnerabilities were high severity, 69 percent were medium severity, and 27 percent were low severity.
- Sixty-one percent of vulnerabilities disclosed during this period affected Web applications, down from 66 percent in the second half of 2006.
- Seventy-two percent of vulnerabilities documented in this reporting period were easily exploitable. This is a decrease from 79 percent in the previous reporting period.
- In the first half of 2007, all operating systems except Hewlett Packard® HP-UX® had shorter average patch development times than in the second half of 2006.
- Hewlett-Packard HP-UX had an average patch development time of 112 days in the first half of 2007, the highest of any operating system. Sun had the highest average patch development time in the second half of 2006, with 145 days.
- The average window of exposure for vulnerabilities affecting enterprise vendors was 55 days. This is an increase over the 47-day average in the second half of 2006.
- Symantec documented 39 vulnerabilities in Microsoft® Internet Explorer, 34 in Mozilla browsers, 25 in Apple® Safari™, and seven in Opera. In the second half of 2006, 54 vulnerabilities were disclosed for Internet Explorer, 40 for Mozilla browsers, four for Apple Safari, and four for Opera.
- Apple Safari had an average window of exposure of three days in the first half of 2007, the shortest of any browser reviewed during this period. Mozilla browsers had the shortest average window of exposure in the second half of 2006, two days.
- Symantec documented six zero-day vulnerabilities in the first half of 2007, down from the 12 that were reported during the second half of 2006.
- Ninety-seven vulnerabilities were documented in Oracle®, more than any other database during the first half of 2007. Oracle also had the most database vulnerabilities in the second half of 2006, with 168.
- There were 90 unpatched enterprise vendor vulnerabilities in the first half of 2007, which is down from the 94 documented in the second half of 2006. Microsoft had the most unpatched vulnerabilities of any enterprise vendor during both of these periods.
- In the first half of 2007, Symantec documented 237 vulnerabilities in Web browser plug-ins. This is a significant increase over 74 in the second half of 2006, and 34 in the first half of 2006.
- During the first half of 2007, 89 percent of plug-in vulnerabilities disclosed affected ActiveX® components for Internet Explorer. ActiveX components accounted for 58 percent of plug-in vulnerabilities in the second half of 2006.

- Symantec found that more than 50 percent of medium- and high-severity vulnerabilities patched by operating system vendors affected Web browsers or had other client-side attack vectors during this and the previous reporting period. Apple was the sole exception, with 49 percent of the vulnerabilities examined in the first half of 2007 affecting browsers or having client-side attack vectors.

Malicious code trend highlights

- Of the top ten new malicious code families detected in the first six months of 2007, four were Trojans, three were viruses, one was a worm, and two were worms with a virus component.
- In the first half of 2007, 212,101 new malicious code threats were reported to Symantec. This is a 185 percent increase over the second half of 2006.
- During the first half of 2007, Trojans made up 54 percent of the volume of the top 50 malicious code reports, an increase over the 45 percent reported in the final six months of 2006.
- When measured by potential infections, Trojans accounted for 73 percent of the top 50 malicious code samples, up from 60 percent in the previous period.
- During this period, 43 percent of worm infections were reported in the Europe, Middle East, and Africa (EMEA) region.
- North America accounted for 44 percent of Trojans reported this period.
- Threats to confidential information made up 65 percent of the top 50 potential malicious code samples by potential infection reported to Symantec.
- Threats with keystroke-logging capacity made up 88 percent of confidential information threats during this period, as did threats with remote access capability, such as back doors. This is an increase from 76 percent and 87 percent respectively over the previous period.
- Forty-six percent of malicious code that propagated did so over SMTP, making it the most commonly used propagation mechanism.
- During the first half of 2007, 18 percent of the 1,509 documented malicious code instances exploited vulnerabilities.
- Thirty-five percent of infected computers reported more than one infection in the first half of 2007.
- Eight of the top ten staged downloaders this period were Trojans and two were worms.
- Seven of the top ten downloaded components were Trojans and three were back doors.
- Malicious code that targets online games made up five percent of the top 50 malicious code samples by potential infection.
- Lineage and World of Warcraft were the two most frequently targeted online games in the first half of 2007.

Phishing Highlights

- The Symantec Probe Network detected a total of 196,860 unique phishing messages, an 18 percent increase over the last six months of 2006. This equates to an average of 1,088 unique phishing messages per day for the first half of 2007.
- Symantec blocked over 2.3 billion phishing messages, an increase of 53 percent over the second half of 2006. This means that Symantec blocked an average of roughly 12.5 million phishing emails per day over the first six months of 2007.
- Organizations in the financial services sector accounted for 79 percent of the unique brands that were used in phishing attacks during this period.
- The brands of organizations in the financial services sector were spoofed by 72 percent of all phishing Web sites.
- Fifty-nine percent of all known phishing Web sites were located in the United States, a much higher proportion than in any other country.
- Three phishing toolkits were responsible for 42 percent of all phishing attacks observed by Symantec in the first half of 2007.
- Eighty-six percent of all phishing Web sites were hosted on only 30 percent of IP addresses known to be phishing Web servers.

Spam Highlights

- Between January 1 and June 30, 2007, spam made up 61 percent of all monitored email traffic. This is a slight increase over the last six months of 2006 when 59 percent of email was classified as spam.
- Sixty percent of all spam detected during this period was composed in English, down from 65 percent in the previous reporting period.
- In the first half of 2007, 0.43 percent of all spam email contained malicious code compared to 0.68 percent in the second half of 2006. This means that one out of every 233 spam messages blocked by Symantec Brightmail AntiSpam™ in the current reporting period contained malicious code.
- Spam related to commercial products made up 22 percent of all spam during this period, the most of any category.
- During the first six months of 2007, 47 percent of all spam detected worldwide originated in the United States, compared to 44 percent in the previous period.
- In the first six months of 2007, 10 percent of all spam zombies in the world were located in the United States, more than any other country.
- In the first half of 2007, 27 percent of all spam blocked by Symantec was image spam.

Executive Summary Discussion

This section will discuss selected security metrics from the *Internet Security Threat Report* in greater depth, providing analysis and discussion of the trends indicated by the data. The following metrics will be discussed:

- Malicious activity originating from Fortune 100 companies
- Data breaches that could lead to identity theft
- Underground economy servers
- Bot-infected computers
- Browser plug-in vulnerabilities
- New malicious code threats
- Trojans
- Threats to confidential information
- Malicious code that targets online games
- Phishing
- Spam

Malicious activity originating from Fortune 100 companies

For the first time, in this volume of the *Internet Security Threat Report*, Symantec is evaluating the amount of malicious activity originating from the IP space of computers and networks that are known to belong to Fortune 100 organizations. Briefly, these are the companies that are determined by Fortune magazine to be the 100 highest grossing companies in the world.[2] Symantec has compiled data on numerous malicious activities that were detected originating from the IP address space of these companies.[3] These activities include: bot-infected computers, phishing Web sites, spam zombies, and Internet attacks.

This metric is significant because it indicates the level to which Fortune 100 organizations have been compromised and are being used by attackers as launching pads for malicious activity. This could affect the performance of the company's networks, thereby reducing employee productivity and limiting the ability of customers to access organizational resources. It could also potentially expose proprietary information, which could have serious business ramifications. Finally, attack activity originating from the organization's network could have serious legal consequences for the company.

Between January 1 and June 30, 2007, four percent of malicious activity detected by Symantec originated from the IP address space of Fortune 100 companies. The IP space of Fortune 100 organizations constitutes just over seven percent of the world's active and advertised IP space.[4] Since the proportion of malicious activity originating from Fortune 100 IP space is lower than the proportion of the world's active and advertised IP space that is assigned to these organizations, less attack activity is originating from Fortune 100 companies than other IP spaces. It is likely that security measures put in place on Fortune 100 networks make it difficult for attackers to compromise them, or to use them to launch attack activity. It could also be due to the fact that some Fortune 100 companies may not use all of the IP space allotted to them. Despite this, networks and computers within these organizations are likely enticing targets for attackers.

[2] http://money.cnn.com/magazines/fortune/fortune500/2007
[3] IP addresses for Fortune 100 companies were determined using autonomous system number (ASN) information.
[4] IP addresses used to determine this proportion were derived from autonomous system number (ASN) information.

There are a number of reasons an attacker may specifically target a Fortune 100 company. By initially targeting well known companies such as these, attackers are targeting victims indirectly by first exploiting trusted entities and then using their position on the network of the trusted company to attack the real victims. Computers within a Fortune 100 company offer attackers many benefits not offered by other computers. For instance, a single compromised computer within such an organization could allow an attacker to gain access to other computers within the organization. This could allow the attacker to harvest various types of information, including the organization's customer database, financial activities of the organization, and proprietary technology or software, to name a few.

Fortune 100 companies also present an attractive target for phishers. For example, an attacker could use a compromised Web server within a Fortune 100 retail company to host phishing Web sites that target customers of the company. Since the phishing Web site would actually be on the compromised company's Web server, customers may be unable to identify it as being fraudulent. An attacker could also send phishing emails from a compromised mail server within a Fortune 100 company's network, which would have a similar obfuscating effect.

To maintain secure networks, organizations should employ defense-in-depth strategies,[5] including the deployment of intrusion detection/intrusion prevention systems (IDS/IPS), antivirus and antifraud solutions and a firewall. Users should update antivirus definitions regularly and ensure that all desktop, laptop, and server computers within an organization are updated with all necessary security patches from their respective vendors. Symantec also advises that policies exist that prevent users from viewing, opening, or executing any email attachment unless the attachment is expected and comes from a known and trusted source, and unless the purpose of the attachment is known.

Data breaches that could lead to identity theft

Identity theft is an increasingly prevalent security issue, particularly for organizations that store and manage information that could facilitate identity theft. Compromises that result in the loss of personal data could be quite costly, not only to the people whose identity may be at risk and their respective financial institutions, but also to the organization responsible for collecting the data.

Data breaches that lead to identity theft could damage an organization's reputation, and undermine customer and institutional confidence in the organization. With the implementation of recent legislation in some jurisdictions,[6] organizations could also be held liable for data breaches and losses, which may result in fines or litigation.[7]

In the first half of 2007, the education sector accounted for more data breaches that could lead to identity theft than any other sector, making up 30 percent of the total (figure 1). This is up from the previous period when the education sector accounted for only 22 percent of the total and ranked second.

[5] Defense-in-depth emphasizes multiple, overlapping, and mutually supportive defensive systems to guard against single-point failures in any specific technology or protection methodology. Defense-in-depth should include the deployment of antivirus, firewalls, and intrusion detection systems, among other security measures.
[6] http://www.parliament.the-stationery-office.co.uk/pa/cm199900/cmbills/001/2000001.htm
[7] http://www.fsa.gov.uk/pages/Library/Communication/PR/2007/021.shtml

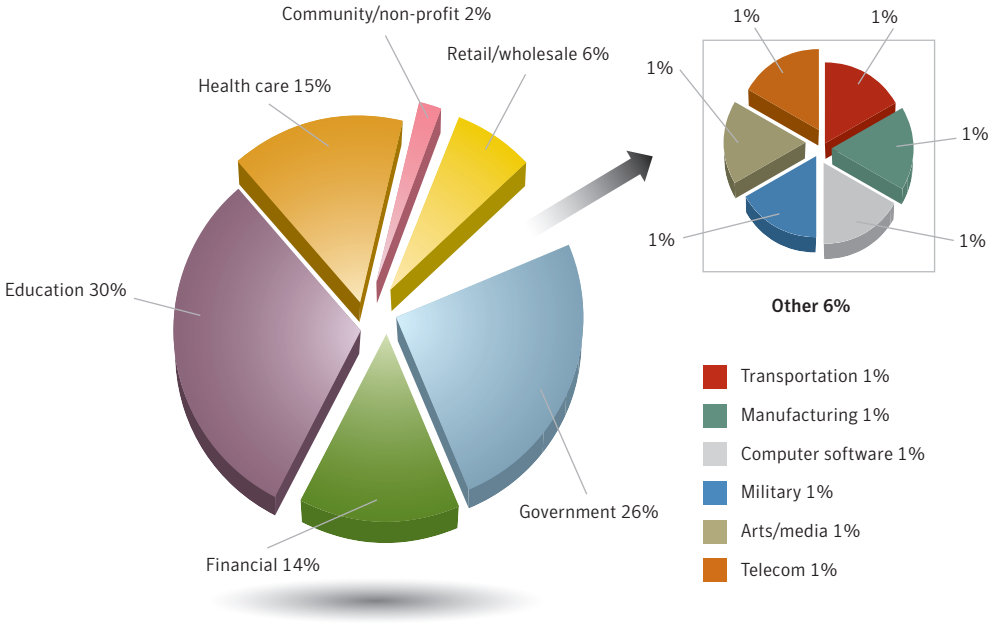

Figure 1. Data breaches that could lead to identity theft by sector
Source: Based on data provided by Attrition.org

Educational organizations store a lot of personal information that could be used for the purposes of identity theft. These organizations—particularly larger universities—often consist of many semi-independent departments in which sensitive personal identification information may be stored in separate locations and be accessible by many people. This increases the opportunities for attackers to gain unauthorized access to this data. Adding to this is the fact that research hospitals, which are considered part of the education sector, store considerable amounts of patients' personal data, including medical information.

During the first half of 2007, the retail/wholesale sector accounted for only six percent of all data breaches that could lead to identity theft, making it the fifth ranked sector during this period. However, the sector was responsible for the largest number of exposed identities, accounting for 85 percent. Breaches in this sector were thus more likely to lead to wide-scale identity theft than any other sector.

The prominence of the retail/wholesale sector was primarily due to the data breach involving the TJX group of retail companies.[8] TJX was a victim of an extensive attack that exposed over 45 million credit and debit card numbers. The number of identities exposed through this breach alone made up over 70 percent of all identities exposed during the period. Due to the nature and extended time span of the compromise, it is likely that these breaches were due to a failure of effective security policies.[9]

In the first half of 2007, the primary cause of data breaches that could facilitate identity theft was the theft or loss of a computer or other medium on which data is stored or transmitted, such as a USB key or a back-up medium. These made up 46 percent of all such data breaches during this period. Theft or loss accounted for 57 percent of all reported breaches in the previous reporting period. Despite this, theft or loss of a computers and storage media only accounted for 11 percent of all identities exposed.

[8] http://www.securityfocus.com/brief/441
[9] http://www.theregister.co.uk/2007/05/04/txj_nonfeasance/

Thus, although theft or loss of computers and computer media is extremely common, it can be considered less likely to lead to wide-scale identity theft than other causes, as it results in relatively fewer exposed identities. This is likely because, in many cases, theft or loss of a computer or computer media is driven not by a desire to steal data, but to steal the hardware itself. A person who steals a laptop is likely driven by the desire to simply sell the laptop for financial gain, and not to harvest the data it may store.

In the first six months of 2007, hacking was the third leading cause of data breaches that could lead to identity theft, accounting for 16 percent of the total. However, it was responsible for 73 percent of identities compromised during the period. A data breach is considered to be caused by hacking if identity theft-related data was exposed by an attacker or attackers by gaining unauthorized access to computers or networks. The prominence of hacking as a cause of compromised identities was largely driven by the TJX breach that was discussed previously in this section.

Because it is responsible for a large number of identities being compromised, hacking is considered one of the causes of data breaches most likely to lead to wide-scale identity theft. This is likely because hacking is more clearly purpose-driven than lost devices or insecure policy. It is an intentional act with a clearly defined purpose—to steal data that can be used for purposes of identity theft or other fraud.

Most breaches that could lead to identity theft are avoidable. In the case of theft or loss, the compromise of data could be averted by encrypting all sensitive data. This would ensure that even if the data is lost or stolen, it would not be accessible to unauthorized third parties. This step should be part of a broader security policy that organizations should develop, implement, and enforce in order to ensure that all sensitive data is protected from unauthorized access.

Organizations can further protect against security breaches that may lead to identity theft by employing defense-in-depth strategies, including the deployment of IDS/IPS solutions, antivirus and antifraud solutions, and a firewall. Antivirus definitions should be updated regularly and all desktop, laptop, and server computers within the organization should be updated with all necessary security patches from their respective vendors.

To help prevent accidental or intentional data leaks, organizations should employ data leakage prevention solutions. Symantec also advises organizations to develop and implement policies that prevent users from viewing, opening, or executing any email attachment unless the attachment is expected and comes from a known and trusted source, and unless the purpose of the attachment is known.

Underground economy servers

Underground economy servers are used by criminals and criminal organizations to sell stolen information, typically for subsequent use in identity theft. This data can include government-issued identification numbers, credit cards, bank cards, personal identification numbers (PINs), user accounts, and email address lists. The emergence of underground economy servers as the *de facto* trading place for illicit information is indicative of the increased professionalization and commercialization of malicious activities over the past several years.

Symantec tracks and assesses underground economy servers across the Internet using proprietary online fraud monitoring tools. For the first time, in this issue of the *Internet Security Threat Report*, Symantec is assessing the types of goods that are most frequently offered for sale on underground economy servers. During the first half of 2007, credit cards were the most frequently advertised item, making up 22 percent of all goods advertised on underground economy servers (table 1).

Rank	Item	Percentage	Range of Prices
1	Credit Cards	22%	$0.50–$5
2	Bank Accounts	21%	$30–$400
3	Email Passwords	8%	$1–$350
4	Mailers	8%	$8–$10
5	Email Addresses	6%	$2/MB–$4/MB
6	Proxies	6%	$0.50–$3
7	Full Identity	6%	$10–$150
8	Scams	6%	$10/week
9	Social Security Numbers	3%	$5–$7
10	Compromised UNIX® Shells	2%	$2–$10

Table 1. Breakdown of goods available for sale on underground economy servers
Source: Symantec Corporation

During the first six months of 2007, Symantec observed 8,011 distinct credit cards being advertised for exchange on underground economy servers. This is only a small proportion of the credit cards sold, however. Typically, users selling credit card information advertise bulk rates and merely give examples of credit card information to attract buyers. Common bulk amounts and rates seen by Symantec during the first six months of 2007 were: 10 credit card numbers for $20 USD; 50 credit card numbers for $70 USD; and 100 credit card numbers for $100 USD.

Symantec also determined that the 85 percent of credit and debit cards advertised for sale on underground economy servers in the first half of 2007 were issued by banks in the United States. This is down slightly from 86 percent in the last six months of 2006.

At the end of 2005, there were approximately 1.3 billion credit cards in circulation in the United States, substantially more than any other country. This likely explains the prominence of US banks in this consideration.[10] Furthermore, the average citizen of the United States has just over four credit cards.[11] If a credit card holder has a large number of credit cards, and uses them all on a regular basis, it is reasonable to assume that monitoring them for illicit use could become difficult.

Identifying fraudulent charges may be even more difficult if they are small or relatively insignificant. For example, small charges may occur when a fraudster attempts to verify whether a card is active by using the stolen card to donate a small amount of money to a charity.[12] If the transaction is successful, the credit card information is then sold or bought. If such a small charge is not identified, the stolen card will likely be used later to commit greater fraud.

[10] http://www.bis.org/publ/cpss78p2.pdf
[11] http://www.bis.org/publ/cpss78p2.pdf
[12] http://www.symantec.com/enterprise/security_response/weblog/2007/07/scammers_make_friends_with_cha.html

The proportion of credit cards advertised matches closely with the market share of each brand of credit card.[13] This implies that the identity-theft community is not specifically targeting any credit card brand.

In order to reduce the likelihood of data breaches that could lead to identity theft, organizations that store personal information should take the necessary steps to protect data transmitted over the Internet or stored on their computers. This should include the development, implementation, and enforcement of secure policy requiring that all sensitive data is encrypted. This would ensure that, even if the computer or medium on which the data were lost or stolen, the data would not be accessible. This step should be part of a broader security policy that organizations should develop and implement in order to ensure that any sensitive data is protected from unauthorized access.

Bot-infected computers

Bots are programs that are covertly installed on a user's machine in order to allow an unauthorized user to control the computer remotely. They allow an attacker to remotely control the targeted system through a communication channel such as IRC. These channels allow the remote attacker to control a large number of compromised computers over a single, reliable channel in a bot network, which can then be used to launch coordinated attacks.

Bots allow for a wide range of functionality and most can be updated to assume new functionality by downloading new code and features. Bots can be used by external attackers to perform DoS attacks against an organization's Web site. Furthermore, bots within an organization's network can be used to attack other organizations' Web sites, which can have serious business and legal consequences. They can be used by attackers to harvest confidential information from compromised computers, which can lead to identity theft. Bots can also be used to distribute spam and phishing attacks, as well as spyware, adware, and misleading applications.

An active bot-infected computer is one that carries out at least one attack per day. This does not have to be continuous; rather, a single computer can be active on a number of different days. Between January 1 and June 30, 2007, Symantec observed an average of 52,771 active bot-infected computers per day (figure 2), a 17 percent decrease from the previous reporting period.

[13] http://www.cardweb.com/cardtrak/pastissues/december2004.html

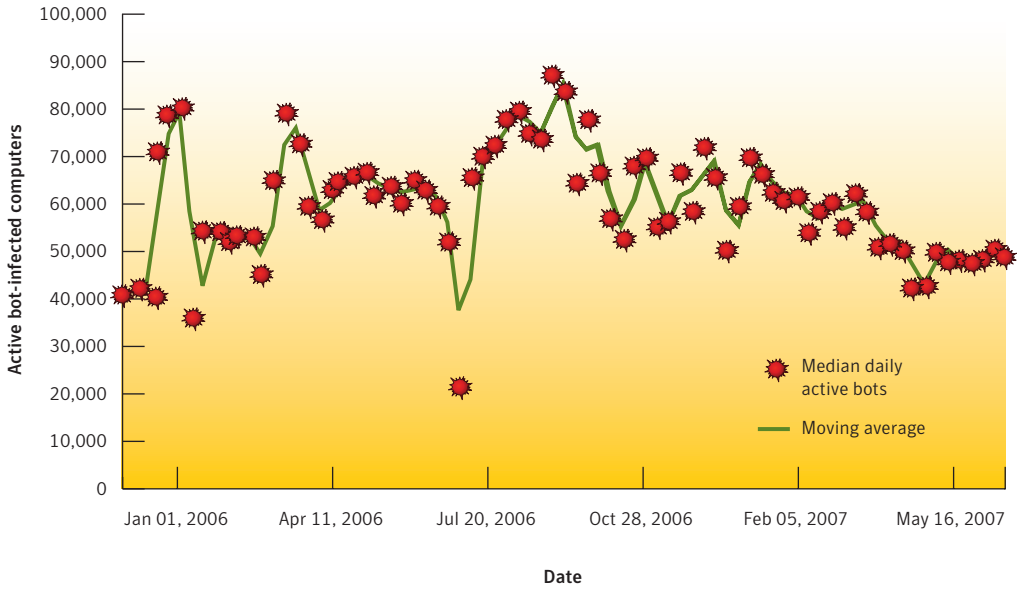

Figure 2. Active bot-infected computers per day
Source: Symantec Corporation

A distinct bot-infected computer is a distinct computer that was active at least once during the period. Symantec also observed 5,029,309 distinct bot-infected computers during this period, a 17 percent decrease from the last six months of 2006.

The decrease in bots observed over the past six months is likely due to a number of reasons, the primary one likely being a change in bot attack methods. As has been discussed in previous volumes of the Symantec *Internet Security Threat Report*, the exploitation of network-based vulnerabilities to spread bots is being slowly abandoned for methods that are more likely to succeed, such as bots that send a mass mailing of themselves.[14] Network-based attacks have been limited somewhat by the introduction of default firewalls in popular operating systems such as Microsoft Windows® XP, as well as an increasing awareness of computer security issues among organizations and computer users. As a result, their use has declined, which has had the effect of limiting the propagation of bots.

Furthermore, law enforcement initiatives targeting bot-networks may also be having some effect. Recently the Federal Bureau of Investigation (FBI) in the United States released information on Operation Bot Roast. This is an ongoing cyber-crime initiative aimed at dismantling bot networks by identifying and arresting bot network owners and taking down the command-and-control servers by which they control their networks.[15] Initiatives such as these will likely result in a reduction in bots for a number of reasons. Firstly, as bot networks are dismantled, less bot activity will be observed. Secondly, as bot network owners become aware of the scrutiny of law enforcement agencies, they are likely to alter their tactics to avoid detection.

The lifespan of a bot is defined as the amount of time that elapses between the first detection of a bot-infected computer until the time that the computer is no longer actively attacking for 30 days, after which time it is assumed to have been disinfected. Gauging the average lifespan of bot-infected computers is important because it allows Symantec to assess how long bot-infected computers are present on a particular network prior to removal.

[14] For instance, please see Symantec *Internet Security Threat Report*, Volume IX (March 2006): http://eval.symantec.com/mktginfo/enterprise/white_papers/ent-whitepaper_symantec_internet_security_threat_report_ix.pdf : p. 30
[15] http://www.fbi.gov/pressrel/pressrel07/botnet061307.htm

During the first six months of 2007, the lifespan of the average bot-infected computer was four days. This is an increase from the previous period, where the average lifespan for a bot-infected computer was three days. The median lifespan of a bot-infected computer during both periods was one day. This indicates that the majority of bot-infected computers are only active for a short period, after which they are identified and disabled, or they are used for activities other than carrying out Internet attacks. The longest lifespan of a bot-infected computer during the period was 3.2 years. However, bots with such long life spans are rare.

The change in the average number of days from three to four from the previous period to the current is likely insignificant. Since the median remained the same, the change in overall average is driven by the longer-lasting bot-infected computers. Given that more time has passed, the age of the longer-lasting bot-infected computers has increased, and so has increased the mean lifespan. Thus, the bot lifespan is holding steady.

It appears that initiatives such as the FBI's Operation Bot Roast, which was discussed previously in this section, are not reducing the lifespan of bot-infected computers. This is likely because the focus of those methods is to eliminate infections and keep infected computers free of bot software, and not necessarily to shorten their effective lives. This is supported by the fact that the number of bot-infected computers has decreased while their lifespan remains steady.

China had the highest number of bot-infected computers during the first half of 2007, accounting for 29 percent of the worldwide total (figure 3), up from 26 percent in the second half of 2006. This continues a trend that was first discussed in the first half of 2005, which saw an increase in bot activity in China during that period.

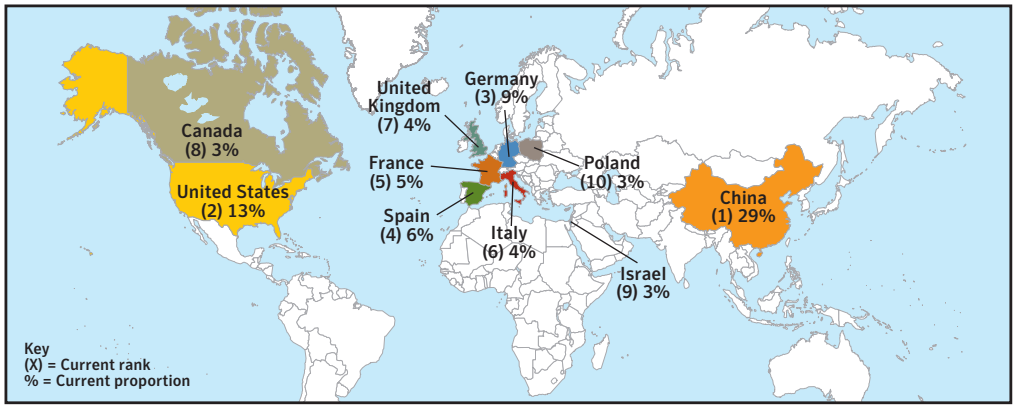

Figure 3. Bot-infected computers by country
Source: Symantec Corporation

Symantec has observed that bots usually infect computers that are connected to high-speed broadband Internet through large Internet service providers (ISPs) and that the expansion of broadband connectivity often facilitates the spread of bots. China's Internet infrastructure is currently expanding rapidly.[16] However, it is worth noting that China's increase in bot-infected computers appears to be slowing. This may be a sign that the security infrastructure as well as awareness is beginning to catch up with Internet user growth.

[16] http://www.vnunet.com/vnunet/news/2163552/china-lead-broadband-world

Command-and-control servers are computers that bot network owners use to relay commands to bot-infected computers on their networks. During the first half of 2007, the United States had the most known command-and-control servers worldwide, accounted for 43 percent of the total. This is a marginal increase over the second half of 2006, when 40 percent of all command-and-control servers were located there.

The high proportion of command-and-control servers in the United States likely indicates that servers there control not only bot networks within the country but elsewhere as well. The high proportion of bot-infected computers and command-and-control servers in the United States is driven by its extensive Internet and technology infrastructure. As of June 2006, more than 58 million broadband Internet users were located there, the highest number in the world.[17]

Browser plug-in vulnerabilities

Browser plug-ins are technologies that run inside the Web browser and extend the browser's features. They can include plug-ins that permit additional multimedia content from Web pages to be rendered in the browser. They also includes execution environments that allow applications to be run inside the browser.

In the first half of 2007, Symantec documented 237 vulnerabilities affecting browser plug-ins (figure 4). Of these, 210 affected ActiveX components, 18 affected the Apple QuickTime® plug-in, four affected the Sun™ Java™ browser plug-in, three affected extensions for Mozilla browsers, and two affected the Adobe Acrobat plug-in. Adobe Flash, Microsoft Windows Media Player, and Opera widgets were not affected by any browser plug-in vulnerabilities during this period.

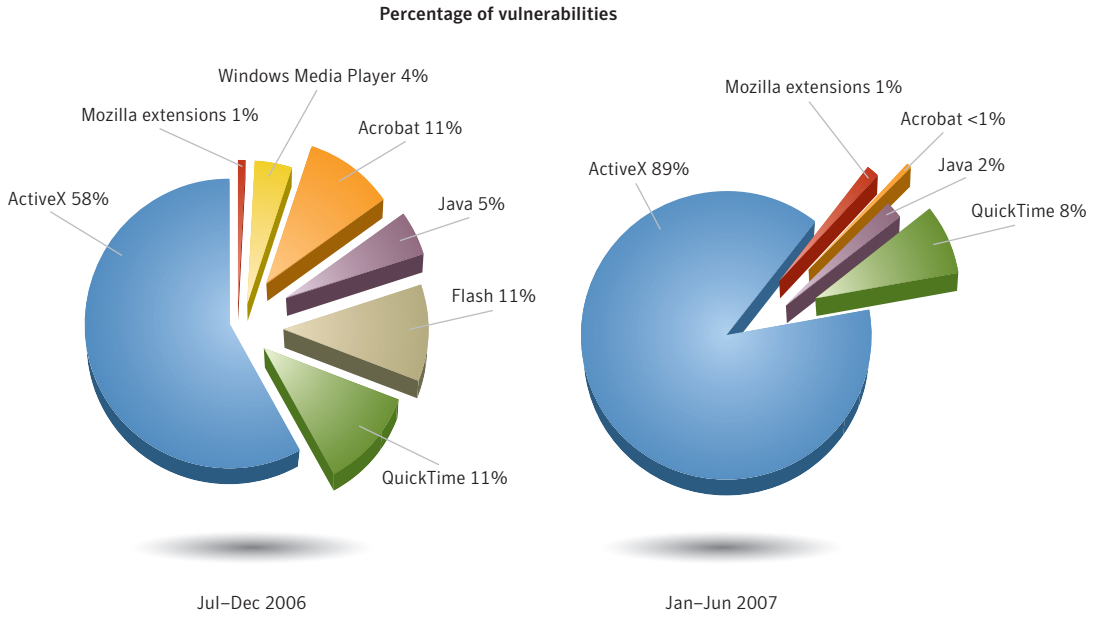

Figure 4. Browser plug-in vulnerabilities
Source: Symantec Corporation

[17] http://www.oecd.org/document/7/0,3343,en_2649_34223_38446855_1_1_1_1,00.html

There were 74 browser plug-in vulnerabilities documented during the second half of 2006. Of those, 43 vulnerabilities affected ActiveX components, eight affected Adobe Flash, eight affected the Apple QuickTime plug-in, seven affected the Adobe Acrobat plug-in, four affected the Sun Java plug-in, three affected Windows Media Player, and one was documented in Mozilla extensions. Opera widgets were not affected by any vulnerabilities in the second half of 2006.

The rise in browser plug-in vulnerabilities is indicative of an increasing focus on client-side vulnerabilities by both security researchers and attackers. The growth corresponds to an increase in the number of vulnerabilities in ActiveX components. This report expands on a previous Symantec study that observed the initial rise in vulnerabilities in ActiveX components.[18] It was determined that the use of fuzzers designed specifically to target insecure ActiveX components has expedited discovery of these vulnerabilities.[19] In addition, it is relatively easy to develop exploits for these types of vulnerabilities due to a high number of previous exploit examples that serve as a template.

These vulnerabilities affect a diverse group of vendors, including Microsoft, enterprise vendors, and smaller vendors. The sheer number of vulnerabilities gives attackers a wide range of potential targets. The installation and execution of ActiveX components is typically transparent to the user, while the removal of such components is not simple for the average end user. As a result, users may not be aware that they are prone to exploitation through vulnerable ActiveX components that have been installed on their computer.

Plug-in vulnerabilities have been the subject of exploit activity in the wild. For example, they were leveraged by many of the exploits employed by the MPack attack framework. In particular, MPack exploits a QuickTime vulnerability,[20] an issue in the WinZip ActiveX component,[21] and various other plug-in vulnerabilities such as the Microsoft WebViewFolderIcon issue.[22]

Client-side attacks have typically originated from questionable sources such as malicious Web sites or spam. As a result, best practices have advised end users to avoid this type of content. However, it appears that attackers are increasingly using legitimate and trusted sites as a basis for attacks. Symantec has observed that MPack includes functionality to serve malicious payloads through legitimate Web sites that have been compromised.[23] MPack is also indicative of a current trend towards multiple staged attacks in which an initial compromise is used to establish a beachhead from which subsequent attacks are launched.

End users and administrators can use a number of measures to protect against the effects of vulnerabilities. IPS technologies can prevent exploitation of some browser plug-in vulnerabilities through signature or behavior-based approaches in addition to address space layout randomization (ASLR). Antivirus software may also aid in protecting organizations from browser plug-in exploits through heuristic signatures.

[18] ActiveX components are a type of COM (Component Object Model) object that may provide a programming interface that is accessible through Internet Explorer. If exposed through Internet Explorer, attackers may exploit latent vulnerabilities in an ActiveX component through malicious HTML content. The study cited is available at: http://www.symantec.com/enterprise/security_response/weblog/2007/01/a_sudden_rise_in_activex_vulne.html
[19] Fuzzing is a security research and quality assurance method that generally entails providing randomly generated inputs in an attempt to discover vulnerabilities and bugs. Fuzzers are programs or scripts that are designed to find vulnerabilities in software code or scripts. They have automated many of the code auditing tasks that security researchers had previously done manually.
[20] http://www.securityfocus.com/bid/21829
[21] http://www.securityfocus.com/bid/21060
[22] http:/www.securityfocus.com/bid/19030
[23] http://www.symantec.com/enterprise/security_response/weblog/2007/06/mpack_the_strange_case_of_the.html

While attacks are likely to originate from Web sites that are trusted as well as those that are not, Web browser security features can help reduce exposure to browser plug-in exploits, as can white-listing. Specifically, administrators and end users should actively maintain a white-list of trusted Web sites, and should disable individual plug-ins and scripting capabilities for all other sites. This will not prevent exploitation attempts from white-listed sites but may aid in preventing exploits from all other sites. Organizations can also implement a white-list policy at the network perimeter to regulate outgoing access by end users.

Trojans

Of the top ten new malicious code families detected in the first six months of 2007, four were Trojans, three were viruses, one was a worm, and two were worms with a virus component. One of the Trojans also had back door capabilities. This indicates that attackers may be moving towards using Trojans as a means of installing malicious code on computers. This is typical of the multiple staged attacks that Symantec is observing with increasing frequency. In these attacks, an initial compromise is not intended to perform malicious activity directly, but is intended to provide a launching point for subsequent, more malicious attack activity.

As Trojans do not propagate, they allow attackers to perform targeted attacks without drawing attention to themselves. Worms, on the other hand, propagate by sending themselves in high volumes of email messages, thereby increasing the likelihood of being noticed by network administrators who can take immediate action. A Trojan that is installed when a user visits a malicious Web site is much more likely to escape notice, as there will be no high-volume traffic associated with it. This increases the Trojan's effectiveness. The longer a threat remains undiscovered in the wild, the more opportunity it has to compromise computers before measures can be taken to protect against it. Furthermore, the longer it can remain resident on a compromised computer, the more confidential information it will be able to steal.

During the first half of 2007, Trojans made up 54 percent of the volume of the top 50 malicious code reports, an increase over the 45 percent reported in the final six months of 2006. While part of this increase can be attributed to the success of the Peacomm Trojan,[24] there were also a wide variety of other Trojans present in the top 50 malicious code reports.

As previously mentioned, Trojans are likely gaining prominence because they generate a low volume of traffic compared to network and mass-mailing worms. As a result, they are less likely to draw the attention of higher-profile threats. Furthermore, malicious code writers may be turning to Trojans because network perimeter defenses and desktop firewalls, neither of which affect Trojans, make it harder for network worms to propagate widely.

The most widely reported new malicious code family during this reporting period was the Peacomm Trojan, also known as the Storm Trojan. This Trojan was spammed in high volumes by the Mixor.Q worm,[25] which prompted Symantec to classify it as a Category 3 threat in January.[26] When Peacomm installs itself on a computer, it attempts to hide itself using rootkit techniques.[27] It also contains a list of other compromised computers that it uses to build an encrypted network of peers, similar to a bot network, although it uses the Overnet peer-to-peer protocol rather than Internet Relay Chat (IRC).[28]

[24] http://www.symantec.com/security_response/writeup.jsp?docid=2007-011917-1403-99
[25] http://www.symantec.com/security_response/writeup.jsp?docid=2006-122917-0740-99
[26] A Category 3 threat is a malicious code sample that is considered a moderate threat. It is either currently spreading among computer users but reasonably harmless and easy to contain, or has not been released into the wild but is potentially dangerous and difficult to contain.
[27] Rootkit techniques are used by malicious code to hide their presence on a compromised computer.
[28] Overnet is a decentralized peer-to-peer file-sharing protocol. It was taken down due to legal action in September 2006, but due to its decentralized nature, clients are still able to function.

Peacomm listens for commands passed through its peer-to-peer (P2P) network and downloads and installs other files, such as the Mespam[29] and Abwiz.F Trojans.[30] This can be of particular concern, since a Trojan like Abwiz.F can send confidential information to the remote attacker and relay spam.

Trojan activity increased from 60 percent of potential infections in the last half of 2006 to 73 percent in the current period. While part of this increase can be attributed to the outbreak of Peacomm in January, there were also a wide variety of other Trojans present in the top 50 malicious code reports. As previously mentioned, Trojans are likely gaining prominence because they generate a low volume of traffic compared to network and mass-mailing worms. As a result, they are less likely to draw the attention of higher-profile threats.

Trojans may also be gaining popularity because they are well suited to meet the objectives of attackers. Trojans are able to perform numerous diverse functions. For example, the Vundo Trojan installs adware on a compromised computer.[31] Variants of the Adclicker Trojan can be used to generate traffic to Web sites in order to increase revenue from banner ads.[32] This practice is frequently referred to as click fraud.[33] Additionally, other Trojans can be used to relay spam email or in phishing attacks. For instance, the Flush Trojan modifies the DNS settings on a compromised computer,[34] which can cause the user's Web browser to be redirected to a phishing site when he or she attempts to connect to an online banking site. The high volume of these Trojans in the top 50 malicious code reports demonstrates the popularity among attackers of utilizing malicious code to generate revenue.

In order to protect against Trojans, administrators and end users should employ defense-in-depth strategies, including the deployment of antivirus software and a personal firewall. Users should update antivirus definitions regularly. They should also ensure that all desktop, laptop, and server computers are updated with all necessary security patches from their software vendors. They should never view, open, or execute any email attachment unless it is expected and comes from a trusted source, and unless the purpose of the attachment is known.

Threats to confidential information

Some malicious code programs are designed specifically to expose confidential information that is stored on an infected computer. These threats may expose sensitive data such as system information, confidential files and documents, or logon credentials. Some malicious code threats, such as back doors, can give a remote attacker complete control over a compromised computer.

Threats to confidential information are a particular concern because of their potential for use in criminal activities. With the widespread use of online shopping and Internet banking, compromises of this nature can result in significant financial loss, particularly if credit card information or banking details are exposed.

Within the enterprise, exposure of confidential information can lead to significant data leakage. If it involves customer-related data—such as credit card information—this can severely undermine customer confidence as well as violate local laws. Sensitive corporate information, including financial details, business plans, and proprietary technologies, could also be leaked from compromised computers.

In the first six months of 2007, threats to confidential information made up 65 percent of potential infections by the top 50 malicious code samples. This is an increase from 53 percent in the second half of 2006.

[29] http://www.symantec.com/security_response/writeup.jsp?docid=2007-020915-2914-99
[30] http://www.symantec.com/security_response/writeup.jsp?docid=2006-032311-1146-99
[31] http://www.symantec.com/security_response/writeup.jsp?docid=2004-112111-3912-99
[32] http://www.symantec.com/security_response/writeup.jsp?docid=2002-091214-5754-99
[33] Click fraud is the act of using illegitimate means, such as a script or program, to imitate the act of a legitimate user clicking on a pay-per-click banner advertisement on a Web page. This act generates revenue for the owner of the page hosting the advertisement. Click fraud is a felony in some jurisdictions.
[34] http://www.symantec.com/security_response/writeup.jsp?docid=2005-030413-5303-99

In this reporting period, remote access threats, such as back door servers, made up 88 percent of confidential information threats (figure 5). They made up 87 percent of confidential information threats in the second half of 2006. Back doors typically require a two-way communication channel between the attacker and the compromised computer in order to access unauthorized information. As such, they can be less efficient than an automated mechanism, such as a keystroke logger. This may indicate why threats that allow remote access only increased marginally this period while other information exposure types increased more significantly.

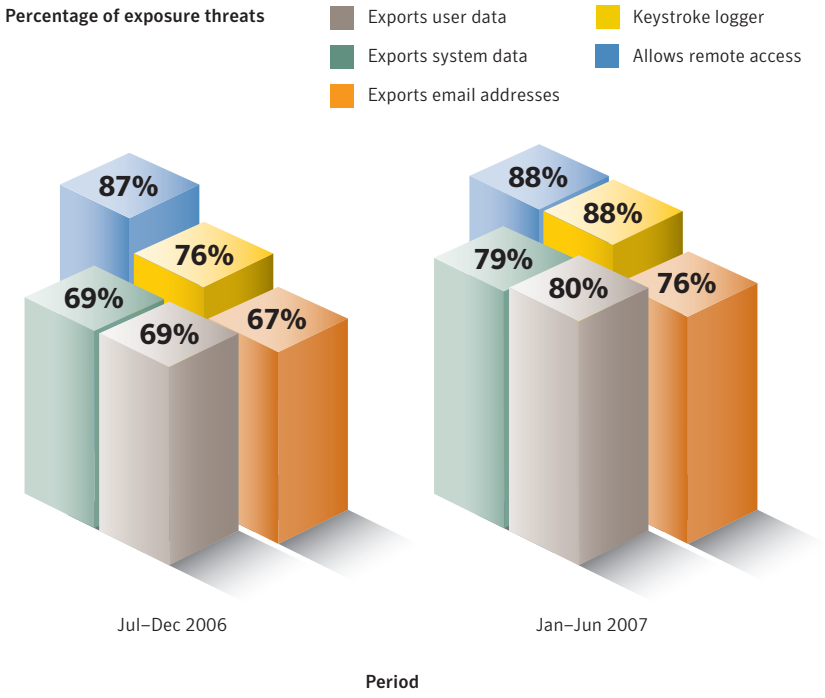

Figure 5. Threats to confidential information by type
Source: Symantec Corporation

Keystroke logging threats made up 88 percent of threats to confidential information, up from 76 percent in the second half of last year. A keystroke logger records keystrokes on a compromised computer and either emails the log to the attacker or uploads it to a Web site under the attacker's control. This makes it easier for the attacker to gather confidential information from a large number of compromised computers than if he or she had to manually connect to back doors installed on various computers.

Malicious code that targets online games

Online gaming is becoming one of the most popular Internet activities. Recently, a study indicated that unique visitors to online gaming sites reached 217 million worldwide. In 2007, the online game market in China alone is expected to grow by 35 percent, where there were 30 million Internet gamers by the end of 2006. Online games often feature goods, such as prizes, that are exchanged by players, often for money. The total annual wealth created within virtual worlds has been placed at approximately 10 billion USD. As such, it is not surprising that attackers appear to be turning their attention to these games.

In the first six months of 2007, five percent of the top 50 malicious code samples reported to Symantec attempted to steal account information for online games. This demonstrates that there is likely considerable financial gain to be made from online gaming accounts, so that attackers are deploying these threats in substantial numbers.

In the first half of 2007, the two most common malicious code sample targeting online games were the Gampass Trojan[35] and the Lineage Trojan.[36] These were also two of the most frequently downloaded components of multistaged downloaders this period. This indicates that attackers see value in targeting online gamers since many of the other top downloaded components are used for more common types of identity theft such as stealing online banking account credentials. Furthermore, the popularity of these staged downloaders illustrates the tendency towards multiple staged attacks that has already been noted in this Executive Summary.

Further reinforcing this notion is the fact that two of the top three malicious code threats targeting online games disable security applications on the compromised computer. This could leave the computer open to other threats even if the user does not participate in any of these online games. Combined with the ability to download other threats, this means that attackers can install a wider range of threats on compromised computers once they have the user's online gaming account information.

Phishing

The Symantec Probe Network blocked over 2.3 billion phishing messages, an increase of 53 percent over the last half of 2007. This means that Symantec blocked an average of roughly 12.5 million phishing emails per day over the first six months of 2007. During this period, Symantec detected a total of 196,860 unique phishing messages, an 18 percent increase over the last six months of 2007. This is an average of 1,088 unique phishing messages per day for the first half of 2007.

For the first time in this volume of the *Internet Security Threat Report*, Symantec is analyzing the usage of automated phishing toolkits in phishing attacks. A phishing toolkit is a set of scripts that allow an attacker to automatically set up phishing Web sites for numerous different brands, including the images and logos associated with those brands. The development and sale of phishing kits is indicative of the increasing professionalization and commercialization in the development and distribution of malicious code and malicious services.

Three phishing toolkits were responsible for 42 percent of all phishing attacks observed by Symantec in the first half of 2007. This shows the high percentage of automation used in phishing attacks. Automation allows attackers to send a high volume of phishing messages that spoof several brands to a large number of recipients with minimal effort. Of the 58 percent of remaining attacks, some may have used phishing toolkits other than the three that are currently known to Symantec, while others used techniques other than toolkits.

[35] http://www.symantec.com/security_response/writeup.jsp?docid=2006-111201-3853-99
[36] http://www.symantec.com/security_response/writeup.jsp?docid=2005-011211-3355-99

Future Watch

This section of the *Internet Security Threat Report* will discuss emerging trends and issues that Symantec believes will become prominent over the next six to twenty-four months. These forecasts are based on emerging research that Symantec has collected during the current reporting period and are speculative in nature. In discussing potential future trends, Symantec hopes to provide organizations and end users with an opportunity to prepare themselves for rapidly evolving and complex security issues. This section will discuss potential security issues associated with the following:

- Malicious code and virtual worlds
- Automated evasion processes—hide and seek for the security generation
- Advanced Web threats—laundering origins through the Web
- Diversification of bot usage

Malicious code and virtual worlds

A persistent virtual world (PVW) is a simulated online environment in which users are able to create personas known as avatars. These avatars are able to interact with each other in a simulated reality environment, 24 hours a day, seven days a week. Second Life is probably the best known example of a PVW.

Virtual worlds often serve as environments in which numerous online users interact in massively multiplayer online games (MMOGs).[37] Popular examples of MMOGs include World of Warcraft and Lineage, both of which allow thousands of players to interact online simultaneously. PVWs and MMOGs are extremely popular, and have been widely adopted in areas like China and South Korea. Symantec believes that as the use of these virtual environments expands, a number of security concerns will emerge.

One simple reason for this is that the main audience of PVWs and MMOGs are early adopters, people who frequently use computers already. As MMOGs become more mainstream, and more commonly played by novice computer users, attack tactics targeting these environments will likely become more effective. The general population (that is, casual players) is probably an audience that attackers will start targeting more.

Many PVWs and MMOGs allow players to conduct real-money transactions (RMTs) in virtual worlds. Players can use credit cards or other payment methods to purchase virtual credits and then exchange those credits with players in other countries, where they may be withdrawn back into local currencies. These RMTs give rise to a de facto international monetary system. There are even exchanges in place for trading (virtual) currency across virtual worlds or different games.[38]

These markets (also referred to as secondary economies) are currently unregulated and are still too small to attract serious attention from law enforcement and securities regulators. Symantec believes that these characteristics could allow criminals to use them for illicit activities. For example, because of the anonymity offered by PVWs, in which all identities are virtual, criminals may be able to launder money through the use of RMTs.

[37] For the purposes of this discussion, MMOGs also include massively multiplayer online role-playing games (MMORPG), which some people consider to be distinct from MMOGs.
[38] http://games.slashdot.org/article.pl?sid=07/06/14/100255&tid=209

To facilitate this, a criminal enterprise could open several thousand MMOG accounts. Each account could be used to trade with other players in the purchase or sale of in-game assets, the funds from which would ultimately be withdrawn from the accounts in question. Since thousands of accounts may engage in millions of transactions, each with small profits or losses, it would be difficult to trace the true source of the funds when they are withdrawn. These transactions can be conducted worldwide without the oversight that typically accompanies international bank remittances. In fact, in February 2007, China's central bank and finance ministries called upon companies to stop trading QQ coins and virtual currencies, presumably to curb the unregulated exchange of currency.[39]

Furthermore, Sparter has created an inter-game currency trading exchange called Gamer2Gamer that permits players to sell their MMOG wares and currencies.[40] Currently, Blizzard Entertainment's World of Warcraft, Turbine's Lord of the Rings Online, Sony Online Entertainment's EverQuest II, and CCP's EVE Online games are supported. Availability of such platforms will further encourage the use of PVWs and MMOGs by attackers as money laundering vehicles.

Symantec also believes that attackers will use PVWs and MMOGs to trick victims into installing malicious software under the pretense that the software improves functionality in the virtual world. For example, virtual worlds have embraced the concept of scripted bots that serve, entertain, and protect avatars within the virtual environment. This could provide attackers with an opportunity to compromise the environment itself.

Although most MMOGs are designed to be played by players, automated tools can be used to enhance play and avoid some tedious, repetitive activities. The downloading and use of these tools presents an opportunity to attackers to incorporate malicious programs such as keystroke loggers and password and information stealers, which the user may unknowingly install on their computer. Symantec has already observed malicious code that attempts to steal information and passwords from players, such as infostealer.wowcraft.[41] Symantec expects that, as in-game toolkits become more popular and are used by more players, attackers will shift their efforts to infecting in-game extensions.

MMOG players and "residents" of virtual communities may also be targeted by phishers and spammers. For instance, users in these environments may receive emails that claim to be from a game's administrators that direct users to spoofed Web sites that are designed to capture account information, such as the player's username and password. The phisher will thus have access to the legitimate player's account, from which they can then distribute the player's assets to other avatars, or sell the account to another player. Despite this risk, the allure of purchasing an established account, with an existing high playing level and established assets at a relative discount (compared to spending thousands of hours playing the game, gaining that level and accumulating similar assets) continues to entice buyers.

Similar to phishing, Symantec also expects to see an increase in the amount of spam that is sent over in-game channels. Spammers will try to collect character names from Web sites that display the standings of the game, or they may use automated scripts to collect player names. Once spam arrives via in-game communications—which may consist of instant messaging clients that are built into the game environment itself—it could be used to deliver phishing attacks or malicious code, or to direct users to malicious Web sites.

[39] http://online.wsj.com/public/article/SB117519670114653518-dn8gNFq5f7FniF4G8iQ_gbzDKug_20080328.html
[40] http://www.shacknews.com/onearticle.x/47408
[41] http://www.symantec.com/security_response/writeup.jsp?docid=2005-073115-1710-99

Automated evasion processes—hide and seek for the security generation

Current antivirus engines are not solely behavior based. Some detect malicious files using static signatures, which simply involve searching for a unique string in a particular file. Others use dynamic analysis, which requires executing the potentially malicious code in a controlled environment. To develop these signatures, antivirus vendors must first acquire malicious code samples through means such as customer submissions, honey pots, or zoo submissions.[42] The samples must then be analyzed, after which signatures are produced and deployed to customers.

The longer a malicious code writer's newest creation goes undetected, the greater the likelihood it will propagate successfully. As malicious code writers put more effort into their creations, the need to evade detection increases. As a result, they have developed numerous evasion mechanisms.

Historically, polymorphism[43] and metamorphism,[44] as well as packers,[45] have been used to evade detection, thereby increasing the effective lifetime of malicious code. However, advances in detecting polymorphic and metamorphic threats and in unpacking malicious code have enabled antivirus vendors to produce signatures that are capable of catching most variations. Malicious code authors have thus been forced to adopt new tactics.

Some of the new techniques center on the distribution point, the point where the malicious code is hosted, such as a Web server. With the significant decline of network-based worms over the past several years (as is discussed in the "Malicious Code Trends" section of this report), current malicious code frequently relies on the exploitation of client-side vulnerabilities. These exploits often use the staged downloader model in which an initial Trojan is installed on the machine and then downloads the most up-to-date version of the malicious code from a distribution point.

Symantec has observed malicious code authors employing numerous techniques to protect the Web servers that are used as distribution points. The most basic is to configure a distributing Web server to serve only one copy of the malicious code per IP address, after which it serves up only a benign executable. The purpose of this is to evade detection and acquisition by security companies who would require samples of the original Trojan in order to produce signatures. This delay in the ability of security companies in acquiring samples increases the chances the malicious code will spread further before detection.

This would have two different consequences. On the one hand, computers behind a Web-proxy or a network address translation device are less likely to become infected since all the computers behind one of these devices share a single IP address. On the other hand, a computer security researcher or malicious code analyst trying to investigate the infection will have trouble obtaining a sample. This difficulty occurs because the same technique could be used to deliberately block IP addresses registered to certain organizations such as antivirus vendors, security consultancies or computer emergency response teams. This phenomenon occurred recently during the MPack Trojan incidents.[46] Malicious code distributors can accomplish these aims either through blacklisting of known IP address ranges or programmatically relying on WHOIS data and performing a keyword search.[47] Symantec expects the prevalence of this defense technique to be more widely deployed in the future due to documented success in instances where it has been used previously.

[42] Malicious code that is developed "in the zoo" is developed in a controlled laboratory environment.
[43] A polymorphic virus is one that can change its byte pattern when it replicates, thereby avoiding detection by simple string-scanning antivirus techniques. In essence, polymorphic viruses make changes to their code to avoid detection.
[44] Metamorphic code evolution describes a method used by malicious code writers that allows a piece of malicious code to change itself autonomously.
[45] Run-time packing utilities, also known as run-time packers, are traditionally used to make files smaller. Malicious code writers use them to make antivirus detection more difficult.
[46] http://www.symantec.com/enterprise/security_response/writeup.jsp?docid=2007-052712-1531-99
[47] WHOIS data stores the name of the person or company who registers a domain and owns IP address space.

Another, more worrisome, technique is known as x-morphism. Borrowing from an idea originally presented by IBM, the concept is simple: the distribution point can serve up a different copy of the malicious code to each visitor. In this scenario, the malicious code no longer has to carry its own metamorphic or polymorphic engine. Instead, the server retains the engine. With this approach, the polymorphic and metamorphic methods that are used to change each instance are hidden, thus making it difficult to produce signatures that reliably work on all variants. Another option available to the malicious code distributor is that the remote site can host a copy of the original source code so any x-morphism can occur in the higher-level programming language before compilation, after which compiler optimization can be used to further obfuscate the sample.

Advanced Web threats—laundering origins through the Web

As the number of available Web services increases and as browsers continue to converge on a uniform interpretation standard for scripting languages such as JavaScript, Symantec expects the number of new Web-based threats to continue increasing. One interesting class of threats includes those that circumvent the same origin policy (SOP) in Web browsers.[48]

One concept that lends itself to SOP circumvention is the mash-up. Mash-ups involve a Web service that collects data from other Web services and then aggregates that data into one view. If data collected from two separate origins is "mashed" through an appropriate Web service, then the end user's Web browser receives the two pieces of data through the same web site. As a result, they appear to have the same origin, even though they may originate from two different sources. Therefore, JavaScript code from one of the origins can obtain and modify properties of the data obtained through the second origin after the two pieces of data have been mashed.

Similar functionality can also be provided by non-transparent Web proxies, like Google Translate. Such proxies generally act as a channel that funnels any content a user desires. Because the content is funneled, from the browser's perspective, the content appears as if it originated from the proxy, when really it might have originated elsewhere. This distinction is important since it might lift restrictions associated with the SOP.

For example, Jikto is a tool that leverages such proxies to scan sites for Web vulnerabilities.[49] The site being scanned and the site containing the scanning code are both loaded through the same proxying service. Therefore, from the Web browser's perspective, they appear to have the same origin, although their actual origins are likely different. As a result, the scanning code can successfully make requests to and read the responses from the site being scanned without being encumbered by the SOP.

Jikto is written entirely in JavaScript so it can run in the user's browser. Any user who visits a page containing the appropriate Jikto source will inadvertently perform a vulnerability scan on a different Web site. That site's Web logs will trace back to the user, and not necessarily to the Web server on which the Jikto source was located. Therefore, since the vulnerability scan is actually being performed by an end user, the attacker's location will be effectively hidden.

Symantec expects that research will continue into novel techniques for SOP circumvention. It is still unclear whether the vulnerabilities found will be exploited in the wild on a wide-scale basis.

[48] The same origin policy dictates that a document or script loaded from one origin (defined with respect to the domain, protocol, and port number) cannot access or modify a document obtained from a different origin. Note that a document or script from one origin can issue a request for a document or script from another origin; however, the first document or script cannot actually read the contents of the other document or script.
[49] http://news.com.com/2100-1002_3-6169034.html

Diversification of bot usage

Bots are programs that are covertly installed on a user's machine in order to allow an unauthorized user to control the computer remotely. They allow an attacker to remotely control the targeted system through a communication channel such as IRC. These channels allow the remote attacker to control a large number of compromised computers over a single, reliable channel in a bot network, which can then be used to launch coordinated attacks.

Bots allow for a wide range of functionality and most can be updated to assume new capabilities by downloading new code and features. They can be used by external attackers to perform DoS attacks against an organization's Web site. Furthermore, bots within an organization's network can be used to attack other organizations' Web sites, which can have serious business and legal consequences. Bots can be used by attackers to harvest confidential information from compromised computers, which can lead to identity theft. They can also be used to distribute spam and phishing attacks, as well as spyware, adware, and misleading applications.

Bots tend to be "early adopters" of new functionality because, due to their design, they can easily incorporate new code across widely dispersed bot networks. As such, they can be used as test environments, deploying new malicious functionalities on a variety of targets before making widespread use of them. Because of this capability, Symantec believes that bots and bot networks will likely be used in an increasingly diverse number of ways in the near future.

For instance, bots may be used in client-side phishing attacks against the legitimate owner or users of an infected computer. Malicious code on an infected computer could be used to mimic the legitimate Web site of an organization whose brand is being used in the phishing attack. As a result, the intended victim could be tricked into disclosing personal identity information, which could subsequently be used in fraudulent activity. This approach allows phishers to bypass some traditional phishing protection mechanisms. Further, a phisher using this technique would not have to rely on a Web site that could be taken down if detected.

In another example, bots can give attackers specific access to infected computers that attackers can then use to their advantage. Bot owners may extract location-identifying information such as domain names from infected computers and subsequently advertise that they control a computer within a specific organization. Parties with interest in the targeted organization might pay for the use of the compromised computer to gather information or to conduct attacks. This approach could greatly increase the risk a bot infection poses to an organization.

In a final example of possible new malicious functionality, bots may be used to artificially increase apparent traffic to certain Web sites. In a twist on the traditional concept of click fraud, bots may be used to hijack browsers, steering them toward sites that allow users to submit and vote upon or recommend Web sites. The idea behind this is to falsely improve search engine ratings, giving the impression of high traffic to a particular site, thereby driving traffic to that site. This could be then used to generate advertising revenue or to serve malicious code, which can then be used in subsequent fraudulent activities.

Attack Trends

This section of the *Internet Security Threat Report* will provide an analysis of attack activity, data breaches that could lead to identity theft, and the trade of illicit information that Symantec observed between January 1 and June 30, 2007. Attacks are defined as any malicious activity carried out over a network that has been detected by an intrusion detection system (IDS) or firewall.

The Symantec Global Intelligence Network, which includes Symantec DeepSight™ Threat Management System and Symantec Managed Security Services, tracks attack activity across the entire Internet. It consists of over 40,000 sensors monitoring network activity in over 180 countries.

Symantec also uses proprietary technologies to monitor bot command-and-control servers and underground economy servers across the Internet. Additionally, Symantec uses publicly available information to assess data breaches that could lead to identity theft.[50] These resources combine to give Symantec an unparalleled ability to identify, investigate, and respond to emerging threats. This discussion will be based on data provided by all of these sources.

Attack Trends Highlights

The following section will offer a brief summary of some of the attack trends that Symantec observed during this period based on data provided by the sources listed above. Following this overview, the *Internet Security Threat Report* will discuss selected metrics in greater depth, providing analysis and discussion of the trends indicated by the data.

- The United States was the country targeted by the most DoS attacks, accounting for 61 percent of the worldwide total in the first half of 2007.
- The United States was the top country of attack origin in the first six months of 2007, accounting for 25 percent of the worldwide attack activity.
- During this period, the United States accounted for 30 percent of all malicious activity, more than any other country.
- Israel was the country with the most malicious activity per Internet user in the first six months of 2007, followed by Canada and the United States.
- Four percent of all malicious activity detected during the first six months of 2007 originated from IP space registered to Fortune 100 companies.
- The education sector accounted for 30 percent of data breaches that could lead to identity theft during this period, more than any other sector.
- Theft or loss of computer or other data-storage medium made up 46 percent of all data breaches that could lead to identity theft during this period.
- The United States was the top country for underground economy servers, accounting for 64 percent of the total known to Symantec.

[50] Data is made available by Attrition.org, a non-profit computer-security related organization: http://www.attrition.org.

- Credit cards were the most common commodity advertised on underground economy servers known to Symantec, accounting for 22 percent of all items.
- Eighty-five percent of credit cards advertised for sale on underground economy servers known to Symantec were issued by banks in the United States.
- Symantec observed an average of 52,771 active bot-infected computers per day in the first half of 2007, a 17 percent decrease from the previous period.
- China had 29 percent of the world's bot-infected computers, more than any other country.
- The United States had the highest number of bot command-and-control servers, accounting for 43 percent of the worldwide total.
- Beijing was the city with the most bot-infected computers, accounting for seven percent of the worldwide total.
- The average lifespan of a bot-infected computer during the first six months of 2007 was four days, up from three days in the second half of 2006.
- Home users were the most highly targeted sector, accounting for 95 percent of all targeted attacks.

Attack Trends Discussion

This section will discuss selected "Attack Trends" metrics in greater depth, providing analysis and discussion of the trends indicated by the data. The following metrics will be discussed:

- Malicious activity by country
- Malicious activity by country per Internet user
- Malicious activity originating from Fortune 100 companies
- Data breaches that could lead to identity theft by sector
- Data breaches that could lead to identity theft by cause
- Underground economy servers by location
- Underground economy servers—credit cards
- Underground economy servers—goods available for sale
- Bot-infected computers
- Lifespan of bot-infected computers
- Bot-infected computers by country

Malicious activity by country

This metric will assess the countries in which the highest amount of malicious activity takes place or originates. To determine this, Symantec has compiled geographical data on numerous malicious activities, namely: bot-infected computers, bot command-and-control servers, phishing Web sites, malicious code reports, spam zombies, and Internet attacks. In addition to data gathered from the Global Intelligence Network, this metric is based on data gathered from the other sources mentioned in the introduction to this report.

To determine the amount of Internet-wide malicious activity that originated in each country, Symantec calculated the mean average of the proportions of all of the aforementioned activities that originated in each country. This average was taken to represent the proportion of overall malicious activity that originated in the country in question and was used to rank each country. This section will discuss those findings.

Between January 1 and June 30, 2007, the United States was the top country for malicious activity, making up 30 percent of worldwide malicious activity (table 2). This represents a minimal change from the second half of 2006, when the United States was also the highest ranked country, accounting for 31 percent of the world's malicious activity. For each of the malicious activities taken into account for this measurement, the United States ranked number one by a large margin with the exception of bot-infected computers. It ranked second for that criteria behind only China.

Overall Rank	Previous Rank	Country	Overall Proportion	Previous Overall Proportion	Malicious Code Rank	Spam Zombies Rank	Command-and-Control Server Rank	Phishing Web sites	Bot Rank	Attack Rank
1	1	United States	30%	31%	1	1	1	1	2	1
2	2	China	10%	10%	2	3	5	18	1	2
3	3	Germany	7%	7%	7	2	2	2	3	3
4	5	United Kingdom	4%	4%	3	15	6	3	7	5
5	4	France	4%	4%	9	7	12	6	5	4
6	7	Canada	4%	3%	6	31	3	7	8	7
7	8	Spain	3%	3%	10	10	22	13	4	6
8	10	Italy	3%	3%	5	6	8	12	6	8
9	6	South Korea	3%	4%	26	8	4	10	13	12
10	11	Japan	2%	2%	4	20	13	8	16	10

Table 2. Malicious activity by country
Source: Symantec Corporation

It is not surprising that the United States was the site of the most malicious activity, as 18 percent of the world's Internet users are located there, more than any other country.[51] Furthermore, it has a well established and relatively long-standing Internet infrastructure. As a result, not only are there a lot of attackers there, but they have had a long time to understand the technologies and to hone their skills. Attackers in countries that have less well established traditions of Internet usage or that are still experiencing rapid growth in their Internet infrastructure may not have the same level of user sophistication.

In previous versions of the *Internet Security Threat Report*, Symantec has argued that as Internet infrastructure becomes established, network and end user security should improve. As Internet users become more sophisticated, so does their knowledge of computer security issues overall. However, the prominence of the United States in this discussion, and the attendant level of malicious activity originating there, indicates that this is not always the case. This is likely because attackers are constantly adapting their attacks to circumvent effective security measures, meaning that even users with a high degree of computer security awareness may be at risk of new attack tactics. Given these considerations, and the country's consistently high ranking in each of the high attack categories, the United States will likely remain number one for malicious activity for some time because of this.

[51] http://www.internetworldstats.com/stats14.htm

China had the second highest amount of malicious activity during the first six months of 2007, accounting for 10 percent of malicious activity detected worldwide, the same rank and percentage as in the previous reporting period. China has the second highest number of Internet users in the world, surpassed only by the United States.[52] However, users in China spend more time online, on average, than those in the United States.[53]

While China ranked highly overall in most of the contributing criteria, it ranked only eighteenth in the world for phishing Web sites. The relatively low ranking of phishing Web sites in China may be linked to the strict regulation of Web sites by the Chinese government,[54] which is enforced through Internet filtering tools on every level from ISPs to Internet cafes.[55]

China also ranked only fifth for bot command-and-control servers, despite the fact that it ranked number one for bot-infected computers. This discrepancy in numbers may indicate that bot-infected computers in China are being controlled by command-and-control servers outside of China. Since the United States has the highest number of command-and-control servers by a large margin, it is likely that bot-network owners in that country are using bot-infected computers in China to conduct attack activity.[56] Thus, some malicious activity attributed to China may not be the result of attackers located there, although the same caveat would also apply to malicious activity originating in other countries as well.

In the first six months of 2007, Germany was the third ranked country for malicious activity. Seven percent of all Internet-wide malicious activity originated there during this period, the same percentage as the second half of 2006 when it was also the third ranked country in this metric. Like both China and the United States, Germany has a well established Internet infrastructure. Furthermore, it has the fourth highest number of Internet users in the world, boasting five percent of the total.[57]

Germany ranks highly in spam zombies, phishing Web sites, bot-infected computers, and command-and-control servers. These activities are often associated with bot networks. As a result, it is likely that bot-networks are prominent in Germany, which would contribute to the high amount of malicious activity originating there.

On a global scale the distribution of the world's malicious activity seems to be relatively static. It appears that a country that is established as a frequent source of malicious activity tends to remain so. This seems to suggest that once an attack infrastructure is established in a country, it becomes entrenched and difficult to remove. Although malicious tools and methods may change, the proportion of malicious activity that originates within a country tends to remain relatively static. This is likely to remain the case until new and more effective measures are taken by countries to reduce the malicious activity originating from their networks.

There are a number of measures that enterprises, administrators, and end users can take to protect against malicious activity. To prevent bot infections, Symantec recommends that ISPs perform both ingress and egress filtering to block known bot traffic.[58] ISPs should also filter out potentially malicious email attachments to reduce exposure to enterprises and end users.

[52] http://www.internetworldstats.com/stats3.htm
[53] http://www.forbes.com/2006/03/31/china-internet-usage-cx_nwp_0403china.html
[54] http://www.cbsnews.com/stories/2002/12/03/tech/main531567.shtml
[55] http://news.bbc.co.uk/2/hi/business/2264508.stm
[56] It should be noted that the location of the command-and-control server does not necessarily correspond to the location of the bot-network owner.
[57] http://www.internetworldstats.com/stats4.htm
[58] Ingress traffic refers to traffic that is coming into a network from the Internet or another network. Egress traffic refers to traffic that is leaving a network, bound for the Internet or another network.

Enterprises should monitor all network-connected computers for signs of malicious activity, ensuring that any infected computers are removed from the network and disinfected as soon as possible. They should also ensure that all antivirus definitions are updated regularly. As compromised computers can be a threat to other systems, Symantec also recommends that enterprises notify their ISPs of any potentially malicious activity.

Organizations should also perform filtering on outgoing network traffic, ensuring that malicious activity and unauthorized communications are not taking place. They should also create and enforce policies that identify and restrict applications that can access the network.

End users should employ defense-in-depth strategies, including the deployment of antivirus software and a firewall. Users should update antivirus definitions regularly and ensure that all desktop, laptop, and server computers are updated with all necessary security patches from their operating system vendor. Symantec also advises that users never view, open, or execute any email attachment unless the attachment is expected and comes from a known and trusted source, and unless the purpose of the attachment is known.

Malicious activity by country per Internet user

Having assessed the top countries by malicious activity, Symantec has also evaluated the top 25 of these countries according to the number of Internet users located there. This measure is intended to remove the bias of high numbers of Internet users from the consideration of the "Malicious activity by country" metric.

In order to determine this, Symantec divided the amount of malicious activity originating in each of the top 25 countries by the number of Internet users who are located in that country. The proportion assigned to each country in this discussion thus equates to the proportion of malicious activity that could be attributed to a single, or average, Internet user in that country. The percentage of malicious activity that would be carried out by each person is the amount assigned to each country in the discussion below.

During the first six months of 2007, Israel was the most highly ranked country for malicious activity per Internet user. If one person from each of the top 25 countries were assessed as a representation of their country's Internet users, the average user in Israel would carry out 11 percent of the group's malicious activity (figure 6). This is a small increase from nine percent in the previous period.

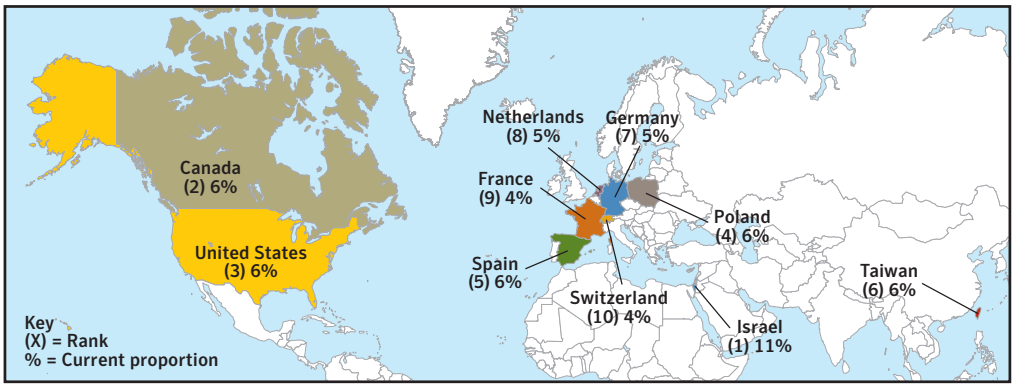

Figure 6. Malicious activity by country per Internet user
Source: Symantec Corporation

This increase was likely due to a higher proportion of bot-infected computers, bot command-and-control servers, and spam zombies located in Israel during this reporting period. This reflects the fact that bots are gaining prominence in Israel; in fact, the number of bot-infected computers located there increased by 15 percent between the second half of 2006 and the end of June 2007.

The prominence of Israel in this metric is likely influenced by the amount of time computer users there spend online. According to a survey released in January 2007, users in Israel spend the second highest number of hours online, on average, less than only users in Canada, which ranked second in this metric.[59] The longer computers are online, the greater the opportunity for attackers to compromise them, particularly through potential vulnerabilities in Internet-based services such as RPC-DCOM and/or client-side applications such as Web browsers.

Furthermore, computer security law enforcement resources in Israel may be insufficient to meet current demands. This prompted a reorganization in 2005 that was intended to create a single information technology authority in the country to deal with computer and Internet crime.[60] As a result of these recent changes, the new security organization may be experiencing difficulties in detecting and eliminating security issues. This is corroborated by the extensive industrial espionage scandal that was uncovered in Israel in 2005.[61]

Finally, ISPs in Israel may not be adequately maintaining secure networks in the country. In 2005, a major Israeli ISP was privatized.[62] The ensuing competition amongst ISPs may have forced those organizations to focus more on expanding their market share than providing the necessary measures for effective computer security.

Canada had the second most malicious activity per Internet user, accounting for six percent of the worldwide total. In the previous reporting period, Canada ranked fifth in this category, with five percent of malicious activity per Internet user. As was discussed previously in this section, Canada had the highest number of hours spent online per person in the first half of 2007. This likely contributes to the country's prominence.

The United States ranked third, accounting for six percent of malicious activity per Internet user. In the second half of 2006, the United States was fourth in this category, but had the same proportion of malicious activity per Internet user. The United States had the fourth highest number of hours spent online per unique Internet user.[63]

The prominence of both Canada and the United states is likely due to the number of hours spent online by the average user and the well established Internet infrastructure in both countries. As was discussed in the "Malicious Activity per Country" metric, the population of Internet users in a country with a well established tradition of usership is more likely to have the skills and experience necessary to conduct sophisticated attack activity. As such it is likely that a higher proportion of the Internet user population would be able to carry out malicious activity, such as creating bot networks, which can then be used for subsequent attack activity.

[59] http://www.websiteoptimization.com/bw/0703
[60] http://www.crime-research.org/news/30.09.2005/1522
[61] http://www.msnbc.msn.com/id/8064757
[62] http://globaltechforum.eiu.com/index.asp?layout=newdebi&country_id=IL&channelid=6&country=Israel&title=Doing+e-business+in+Israel
[63] http://www.websiteoptimization.com/bw/0703/

Both Canada and the United States both ranked lower for malicious activity per Internet user in the second half of 2006 than in the current reporting period. The current increase is primarily driven by a drop in malicious activity per Internet user in Taiwan, which came primarily from a drop in both malicious code infections and bot-infected computers. This is likely due to the fact that bot-infected computers in Taiwan dropped by 46 percent between the second half of 2006 and the end of the first half of 2007.

Malicious activity originating from Fortune 100 companies

For the first time, in this volume of the *Internet Security Threat Report*, Symantec is evaluating the amount of malicious activity originating from the IP space of computers and networks that are known to belong to Fortune 100 organizations. Briefly, these are the companies that are determined by Fortune magazine to be the 100 highest grossing companies in the world.[64] Symantec has compiled data on numerous malicious activities that were detected originating from the IP address space of these companies.[65] These activities include: bot-infected computers, phishing Web sites, spam zombies, and Internet attacks.

This metric is significant because it indicates the level to which Fortune 100 organizations have been compromised and are being used by attackers as launching pads for malicious activity. This could affect the performance of the company's networks, thereby reducing employee productivity and limiting the ability of customers to access organizational resources. It could also potentially expose proprietary information, which could have serious business ramifications. Finally, attack activity originating from the organization's network could have serious legal consequences for the company.

Between January 1 and June 30, 2007, four percent of malicious activity detected by Symantec originated from the IP address space of Fortune 100 companies (figure 7). The IP space of Fortune 100 organizations constitutes just over seven percent of the world's active and advertised IP space.[66] Since the proportion of malicious activity originating from Fortune 100 IP space is lower than the proportion of the world's active and advertised IP space that is assigned to these organizations, less attack activity is originating from Fortune 100 companies than other IP spaces. It is likely that security measures put in place on Fortune 100 networks make it difficult for attackers to compromise them, or to use them to launch attack activity. It could also be due to the fact that some Fortune 100 companies may not use all of the IP space allotted to them. Despite this, networks and computers within these organizations are likely enticing targets for attackers.

[64] http://money.cnn.com/magazines/fortune/fortune500/2007
[65] IP addresses for Fortune 100 companies were determined using autonomous system number (ASN) information.
[66] IP addresses used to determine this proportion were derived from autonomous system number (ASN) information.

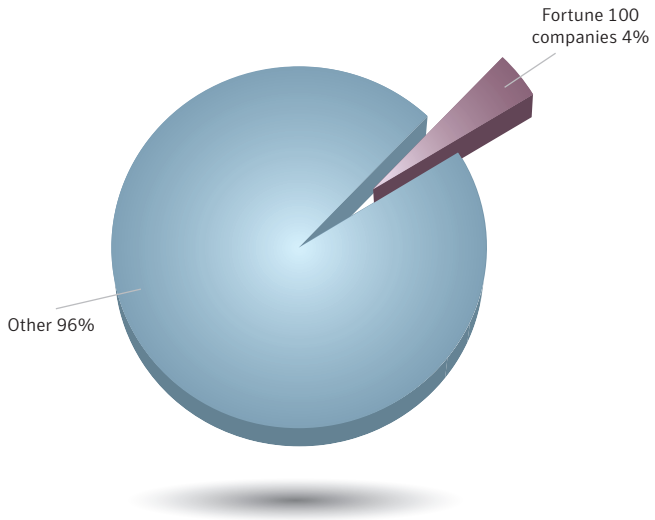

Figure 7. Malicious activity originating from Fortune 100 companies
Source: Symantec Corporation

There are a number of reasons an attacker may specifically target a Fortune 100 company. Computers within a Fortune 100 company offer attackers many benefits not offered by other computers. For instance, a single compromised computer within such an organization could allow an attacker to gain access to other computers within the organization. This could allow the attacker to harvest various types of information, including the organization's customer database, financial activities of the organization, and proprietary technology or software to name a few.

A prominent example of this type of incident is the TJX compromise.[67] TJX is not a Fortune 100 company, but it is a large organization that operates many different retail outlets including T.J. Maxx, T.K. Maxx, Marshalls, and Winners. Attackers compromised the wireless networks of the company, allowing them to steal the personal information of over 45 million customers, including credit card information, which was later used to commit fraud.[68]

Attackers may also be enticed to target Fortune 100 companies in order to gain access to their considerable network resources. Large organizations typically have much higher bandwidth networks than are available to home users. These would give an attacker access to much higher-speed and higher-capacity communications than would attacks against small office and home user computers. This could facilitate a wide variety of attack activity, such as large DoS attacks. It could also potentially allow small attacks to go unnoticed amidst the high volume of standard business traffic.

Fortune 100 companies also present an attractive target for phishers. For example, an attacker could use a compromised Web server within a Fortune 100 retail company to host phishing Web sites that target customers of the company. Since the phishing Web site would actually be on the compromised company's Web server customers may be unable to identify it as being fraudulent. An attacker could send also phishing emails from a compromised mail server within a Fortune 100 company's network, which would have a similar obfuscating effect.

[67] http://www.securityfocus.com/brief/441
[68] http://www.securityfocus.com/news/11438

To maintain secure networks, organizations should employ defense-in-depth strategies, including the deployment of IDS/IPS solutions, antivirus and antifraud solutions and a firewall. Users should update antivirus definitions regularly and ensure that all desktop, laptop, and server computers within an organization are updated with all necessary security patches from their respective vendors. Symantec also advises that policies exist that prevent users from viewing, opening, or executing any email attachment unless the attachment is expected and comes from a known and trusted source, and unless the purpose of the attachment is known.

Data breaches that could lead to identity theft

Identity theft is an increasingly prevalent security issue, particularly for organizations that store and manage information that could facilitate identity theft. Compromises that result in the loss of personal data could be quite costly, not only to the people whose identity may be at risk and their respective financial institutions, but also to the organization responsible for collecting the data. The metrics that follow will assess data breaches that may have exposed information that could lead to identity theft.

Data breaches that lead to identity theft could damage an organization's reputation, and undermine customer and institutional confidence in the organization. With the implementation of recent legislation in some jurisdictions,[69] organizations can also be held liable for data breaches and losses, which may result in fines or litigation.[70] Examples of such legislation include the Health Insurance Probability and Accountability Act (HIPAA),[71] enacted in the United States in 1996, and the Plastic Card Security Act, which was enacted in Minnesota in April 2007.[72] The latter is based on the Payment Card Industry (PCI) Compliance standard.[73]

Data breaches that could lead to identity theft by sector

Using publicly available data,[74] Symantec has determined the sectors that were most often affected by these breaches, as well as the most common causes of data loss. This metric will also explore the number of identities exposed due to these data breaches using the same publicly available data. An identity is considered to be exposed if personal or financial data related to the identity is exposed through the breach.

It should be noted that some sectors may need to comply with more stringent data breach reporting requirements than others. For instance, government organizations are more likely to report data breaches, either due to regulatory obligations or in conjunction with publicly accessible audits and performance reports.[75] Furthermore, organizations that rely on consumer confidence may be less inclined to report such breaches for fear of negative consumer, industry, or market reaction. As a result, sectors that are not required or encouraged to report may be under-represented in this data set.

In the first half of 2007, the education sector accounted for 30 percent of all known data breaches that could lead to identity theft, more than any other sector (figure 8). This is up from the previous period when education accounted for only 22 percent of the total and was the second ranked sector.

[69] http://www.parliament.the-stationery-office.co.uk/pa/cm199900/cmbills/001/2000001.htm
[70] http://www.fsa.gov.uk/pages/Library/Communication/PR/2007/021.shtml
[71] http://www.cms.hhs.gov/HIPAAGenInfo/
[72] http://www.revisor.leg.state.mn.us/bin/bldbill.php?bill=S1574.2.html&session=ls85
[73] Payment Card Industry (PCI) Compliance is a set of security standards that were created by numerous major credit card companies to protect their customers from increasing identity theft and security breaches. For more information, please see: http://www.pcicomplianceguide.org/businesscompliance.html
[74] http://attrition.org/dataloss/
[75] For example, the Fair and Accurate Credit Transactions Act of 2003 (FACTA) of California. For more on this act, please see: http://www.privacyrights.org/fs/fs6a-facta.htm. Another example is the Health Insurance Portability and Accountability Act of 1996. For more information see: http://www.cms.hhs.gov/HIPAAGenInfo/

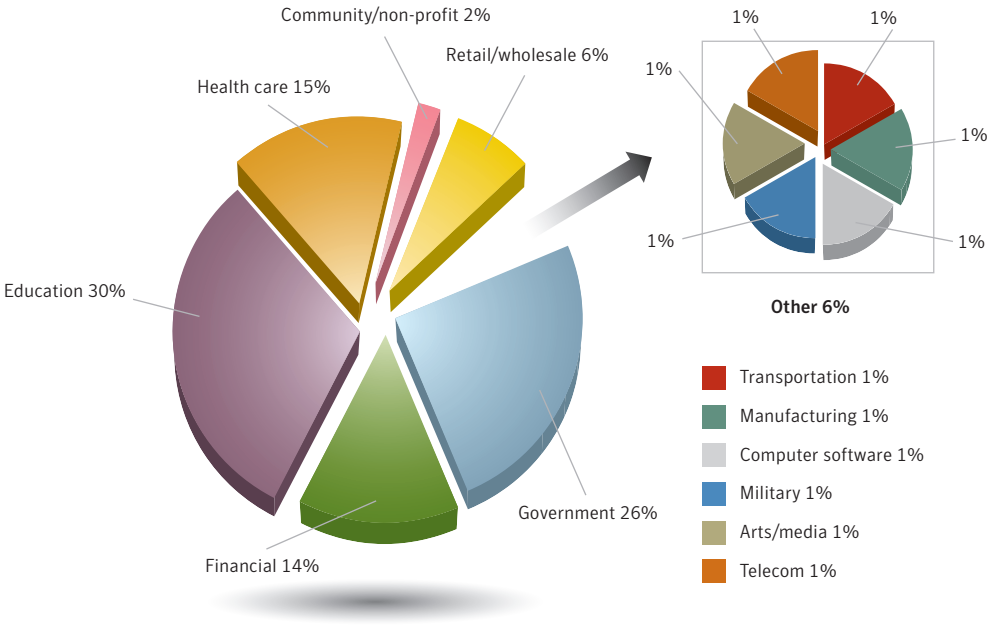

Figure 8. Data breaches that could lead to identity theft by sector
Source: Based on data provided by Attrition.org

Educational organizations store a lot of personal information that could be used for the purposes of identity theft. These organizations—particularly larger universities—often consist of many semi-independent departments in which sensitive personal identification information may be stored in separate locations and be accessible by many people. This increases the opportunities for attackers to gain unauthorized access to this data. Adding to this is the fact that research hospitals, which are considered part of the education sector, store considerable amounts of patients' personal data, including medical information.

In spite of the high number of data breaches that occurred in the education sector during the first six months of 2007, it only accounted for one percent of all identities exposed during the period (figure 9). This is likely because most data breaches within the education sector were caused by theft or loss of computers or data-storage devices. Unlike hacking, in which data breaches can last for an extended period and expose numerous identities, breaches caused by theft or loss can only be opportunistically taken advantage of and cannot provide long term access to large amounts of data.[76] Breaches that occur in the education sector are therefore not as likely to result in wide-scale identity theft because they result in the exposure of relatively few identities.

[76] A data breach is considered to be caused by hacking if identity theft-related data was exposed by an attacker or attackers by gaining unauthorized access to computers or networks.

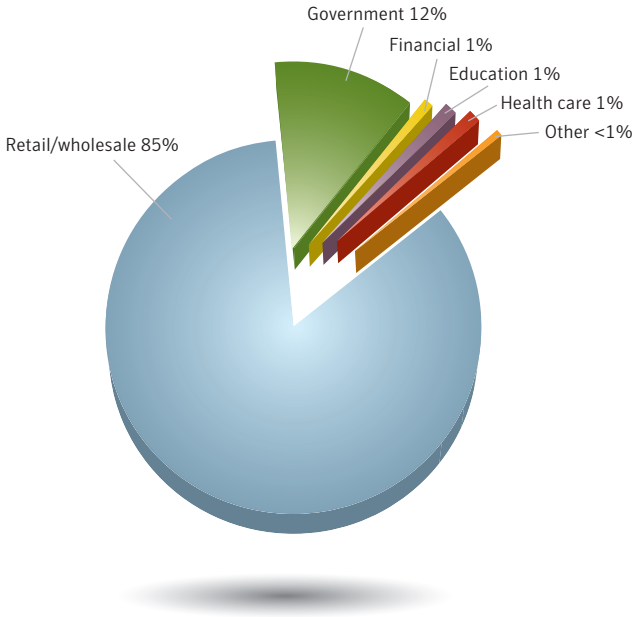

Figure 9. Identities exposed by sector
Source: Based on data provided by Attrition.org

During this reporting period, the government sector accounted for 26 percent of data breaches that could lead to identity theft, making it the second highest sector for this consideration. This sector had the most breaches that could lead to identity theft in the second half of 2006, accounting for 33 percent of the total during that period.

Government organizations, like educational organizations, store a considerable amount of information that could be used for identity theft. Similar to the educational sector, these organizations often consist of numerous semi-independent departments. As a consequence, sensitive personal identification information may be stored in separate locations and be accessible by numerous people. This increases the opportunities for attackers to gain unauthorized access to this data.

The government sector also ranked second for the overall number of identities exposed during the period, accounting for 12 percent of the total. As was the case with the educational sector, the number of identities exposed is relatively small compared to the number of data breaches in this sector. Thus, breaches that occur in the government sector are less likely to result in wide-scale identity theft than those in other sectors.

The health care sector accounted for 15 percent of data breaches that could lead to identity theft in the first half of 2007. Health care ranked fourth in the previous period, accounting for 11 percent of all breaches that could lead to identity theft. The prominence of the health care sector in this metric is likely due to similar factors that influence the prominence of both education and government as outlined previously. Furthermore, health organizations store information related to personal health, which could result in damaging breaches of privacy if viewed by unauthorized people.

The health care sector ranked fifth for the overall number of identities exposed, accounting for just over one percent. So, like both education and government sectors, data breaches within the health care sector resulted in a relatively low number of exposed identities. Thus, breaches in this sector are relatively less likely to result in wide-scale identity theft than those in other sectors because they expose less identity-theft related data.

During the first half of 2007, the retail/wholesale sector accounted for only six percent of all data breaches that could lead to identity theft, making it the fifth ranked sector during this period. However, the sector was responsible for the largest number of exposed identities, accounting for 85 percent. Breaches in this sector were thus far more likely to result in wide-scale identity theft than any other sector. Each data breach would facilitate identity theft to a much greater degree.

The prominence of the retail/wholesale sector was primarily due to the data breach involving the TJX group of retail companies. TJX was a victim of an extensive attack that exposed over 45 million credit and debit card numbers. The number of identities exposed through this breach alone made up over 70 percent of all identities exposed during the period. Due to the nature and extended time span of the compromise, it is likely that these breaches were due to a failure of effective security policies.[77]

Data breaches that could lead to identity theft by cause

In the first half of 2007, the primary cause of data breaches that could facilitate identity theft was the theft or loss of a computer or other medium on which data is stored or transmitted, such as a USB key or a back-up medium (figure 10). These made up 46 percent of all such data breaches during this period. Theft or loss accounted for 57 percent of all reported breaches in the previous reporting period. Despite this, theft or loss of a computers and storage media only accounted for 11 percent of all identities exposed (figure 11). Thus, although theft or loss of computers and computer media is extremely common, it can be considered less likely to result in wide-scale identity theft than other causes, as it results in relatively fewer exposed identities.

This is likely because in many cases, theft or loss of a computer or computer media is driven not by a desire to steal data, but to steal the hardware itself. A person who steals a laptop is likely driven by the desire to simply sell the laptop for financial gain, and not to harvest the data it may store.

[77] http://www.theregister.co.uk/2007/05/04/txj_nonfeasance/

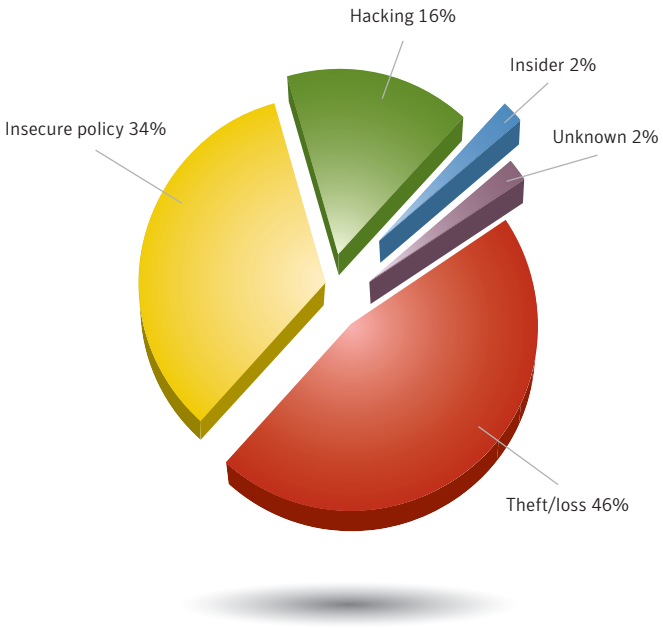

Figure 10. Data breaches that could lead to identity theft by cause
Source: Based on data provided by Attrition.org

The second most common cause of data breaches that could lead to identity theft during this period was insecure policy, which made up 34 percent of all incidents. A data breach is considered to be caused by insecure policy if it can be attributed to a failure to develop, implement, and/or comply with adequate security policy. In the previous period, insecure policy also ranked second, accounting for 27 percent of such data breaches.

In the first half of 2007, insecure policy accounted for only three percent of exposed identities (figure 11). Thus, each breach exposed relatively little personal identity information. This implies that breaches caused by insecure policy are not currently considered particularly likely to result in wide-scale identity theft.

In the first six months of 2007, hacking was the third leading cause of data breaches that could lead to identity theft, accounting for 16 percent of the total. A data breach is considered to be caused by hacking if identity theft-related data was exposed by an attacker or attackers by gaining unauthorized access to computers or networks. During the last six months of 2006, hacking also ranked third, accounting for 11 percent of breaches that could facilitate identity theft.

Hacking was responsible for 73 percent of identities exposed during the period. The prominence of hacking as a cause of exposed identities was largely driven by the TJX breach that was discussed previously in this section. This shows clearly that hacking is the cause of data breaches that is most likely to lead to wide-scale identity theft. This is likely because hacking is more clearly purpose-driven than insecure policy or the loss or theft of devices. It is an intentional act with a clearly defined purpose: to steal data that can be used for purposes of identity theft or other fraud.

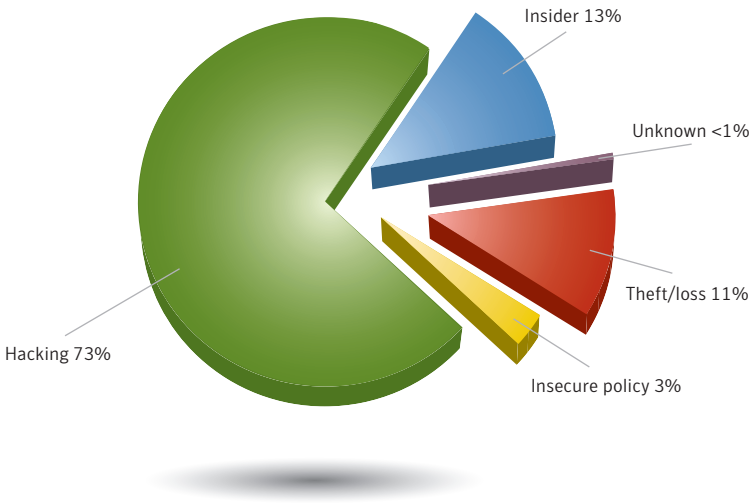

Figure 11. Number of identities exposed by cause
Source: Based on data provided by Attrition.org

Most breaches that could lead to identity theft are avoidable. In the case of theft or loss, the compromise of data could be averted by encrypting all sensitive data. This would ensure that even if the data is lost or stolen, it would not be accessible to unauthorized third parties. This step should be part of a broader security policy that organizations should develop, implement, and enforce in order to ensure that all sensitive data is protected from unauthorized access.

Organizations can further protect against security breaches that may lead to identity theft by employing defense-in-depth strategies, including the deployment of IDS/IPS solutions, antivirus and antifraud solutions, and a firewall. Antivirus definitions should be updated regularly and all desktop, laptop, and server computers within the organization should be updated with all necessary security patches from their respective vendors.

To help prevent accidental or intentional data leaks, organizations should employ data leakage prevention solutions. Symantec also advises organizations to develop and implement policies that prevent users from viewing, opening, or executing any email attachment unless the attachment is expected and comes from a known and trusted source, and unless the purpose of the attachment is known.

Underground economy servers

Underground economy servers are used by criminals and criminal organizations to sell stolen information, typically for subsequent use in identity theft. This data can include government-issued identification numbers (such as Social Security numbers), credit cards, bank cards, personal identification numbers (PINs), user accounts, and email address lists. Symantec tracks and assesses underground economy servers across the Internet using proprietary online fraud monitoring tools.

This discussion will assess underground economy servers in two ways: according to the location of the underground economy server and according to the location of banks that issued credit and debit cards that were being advertised on underground economy servers. It will also look at the different types of items that are being exchanged through underground economy servers as well as the different credit cards and credit card information that is available for sale. It should be noted that this discussion may not necessarily be representative of Internet-wide activity; rather, it is intended as a snapshot of the activity that Symantec monitored during this period.

Underground economy servers by location

During the first six months of 2007, 64 percent of all underground economy servers identified by Symantec were located in the United States, by far the highest total of any country (figure 12). During the last half of 2006, the United States was home to the majority of underground economy servers as well, accounting for 51 percent of the total known to Symantec. The prominence of the United States is likely associated with the relatively high level of malicious activity there, as was discussed previously in this report. This is likely influenced most strongly by the fact that the United States has the highest number of Internet users in the world.[78]

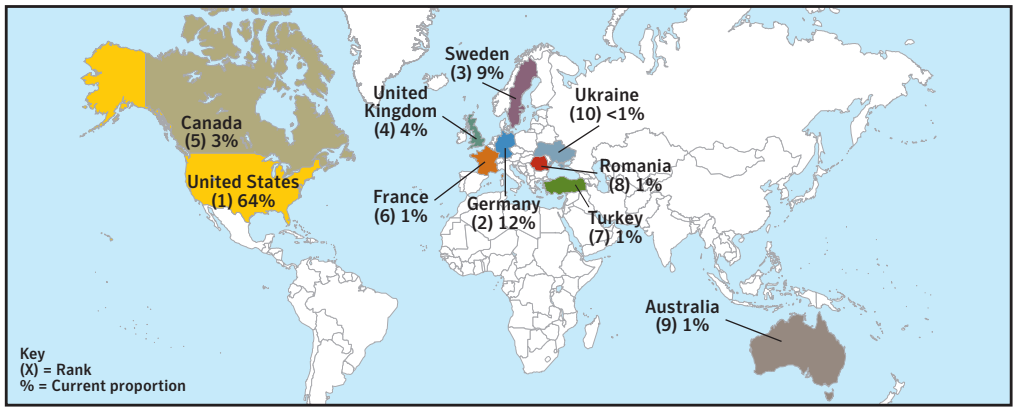

Figure 12. Location of underground economy servers
Source: Symantec Corporation

Germany had the second most economy servers during the first half of 2007, accounting for 12 percent of the worldwide total. In the previous reporting period, Germany ranked fourth, accounting for six percent. Sweden ranked third, accounting for nine percent of worldwide underground economy servers. During the last half of 2006 Sweden ranked second and accounted for 15 percent of all economy servers observed by Symantec.

For each of the top ten countries, the proportion of underground economy servers changed considerably from the previous period. This can be attributed to the nature of these servers, which are often hosted as channels on public IRC servers. Once a fraud-related IRC channel becomes popular, it is often either shut down by the IRC server administrators or abandoned by its users due to legal liability and the increased

[78] http://internetworldstats.com/top20.htm

possibility of being caught. As such, the location of an underground economy server is primarily driven by convenience. Furthermore, the geographic location of the server is typically not of any consequence to those involved; users of underground economy servers do most of their business electronically so they have no geographical restrictions.

Underground economy servers—credit cards

During the first six months of 2007, Symantec observed 8,011 distinct credit cards being advertised for exchange on underground economy servers. This is only a small proportion of the credit cards sold, however. Typically, users selling credit card information advertise bulk rates and merely give examples of credit card information to attract buyers. Common bulk amounts and rates seen by Symantec during the first six months of 2007 were: 10 credit card numbers for $20 USD; 50 credit card numbers for $70 USD; and 100 credit card numbers for $100 USD.

Symantec also determined that the 85 percent of credit and debit cards advertised for sale on underground economy servers in the first half of 2007 were issued by banks in the United States (figure 13). This is down slightly from 86 percent in the last six months of 2006.

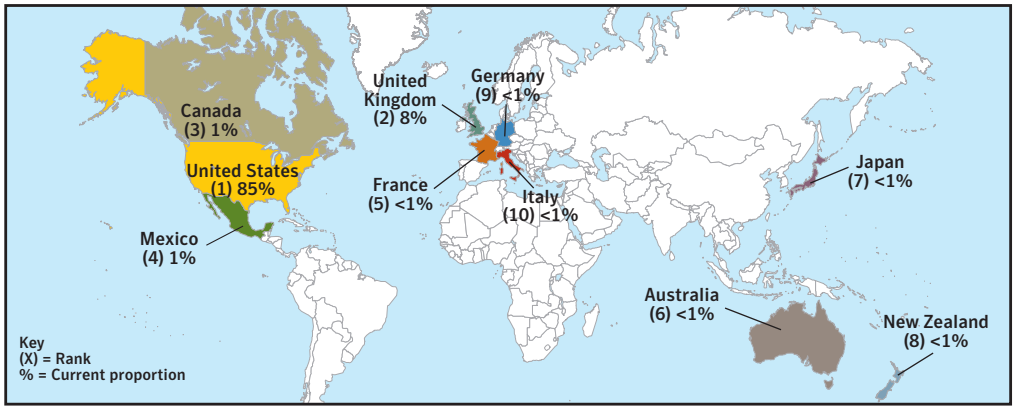

Figure 13. Location of banks whose cards were sold on underground economy servers
Source: Symantec Corporation

At the end of 2005, there were approximately 1.3 billion credit cards in circulation in the United States, substantially more than any other country. This likely explains the prominence of US banks in this consideration.[79] Furthermore, the average citizen of the United States has just over four credit cards.[80] If a credit card holder has a large number of credit cards, and uses them all on a regular basis, it is reasonable to assume that monitoring them for illicit use could become difficult.

Identifying fraudulent charges may be even more difficult if they are small or relatively insignificant. For example, small charges may occur when a fraudster attempts to verify whether a card is active by using the stolen card to donate a small amount of money to a charity.[81] If the transaction is successful, the credit card information is then sold or bought. If such a small charge is not identified, the stolen card will likely be used later to commit greater fraud.

[79] http://www.bis.org/publ/cpss78p2.pdf
[80] http://www.bis.org/publ/cpss78p2.pdf
[81] http://www.symantec.com/enterprise/security_response/weblog/2007/07/scammers_make_friends_with_cha.html

During the first six months of 2007, eight percent of all credit and debit cards advertised on underground economy servers were issued by banks in the United Kingdom, making it the second ranked country, albeit well behind the United States. With just under 70 million credit cards in circulation in the United Kingdom, just over five percent of the number circulating in the United States, the position of the former relative to the latter is not surprising.[82]

Canada ranked third, accounting for one percent of all credit and debit cards advertised on underground economy servers, the same rank and percentage as in the previous six-month period. Canada had just over 60 million credit cards in circulation in 2005.[83] The high number of cards likely has an influence on Canada's ranking. However, Canadian credit cards may be less desirable because criminals using stolen cards may have more trouble using them outside of Canada because of the credit card monitoring practices of Canadian banks.

The proportion of credit cards advertised matches closely with their respective market share.[84] This implies that the identity-theft community is not specifically targeting any credit card brand.

Underground economy servers—goods available for sale

For the first time, in this issue of the *Internet Security Threat Report*, Symantec is assessing the types of goods that are most frequently offered for sale on underground economy servers. During the first half of 2007, credit cards were the most frequently advertised item, making up 22 percent of all goods (table 3).

Rank	Item	Percentage	Range of Prices
1	Credit Cards	22%	$0.50–$5
2	Bank Accounts	21%	$30–$400
3	Email Passwords	8%	$1–$350
4	Mailers	8%	$8–$10
5	Email Addresses	6%	$2/MB–$4/MB
6	Proxies	6%	$0.50–$3
7	Full Identity	6%	$10–$150
8	Scams	6%	$10/week
9	Social Security Numbers	3%	$5–$7
10	Compromised UNIX Shells	2%	$2–$10

Table 3. Breakdown of goods available for sale on underground economy servers
Source: Symantec Corporation

Bank account credentials, including account numbers and authentication information, were the second most commonly advertised item on underground economy servers during the period, accounting for 21 percent of all advertised goods. The advertised price for bank account credentials varied widely, ranging between $30 and $400 USD, and was dependent on the funds available in the account. Bank accounts that included higher balances were worth considerably more. Furthermore, bank account information that included personal information of the victim was more highly valued.

[82] http://www.bis.org/publ/cpss78p2.pdf
[83] http://www.bis.org/publ/cpss78p2.pdf
[84] http://www.cardweb.com/cardtrak/pastissues/december2004.html

Email passwords were the third most common item advertised for sale, making up eight percent of all advertised goods. Email passwords allow access to an email account and are typically used for sending spam. They can also be used to recover a user's passwords from various Web sites that will email password-reset information to the user's email account. The prices for advertised email passwords ranged between $1 USD and $350 USD, depending on whether the account had been used for spamming previously. Furthermore, the value of the account was also based on the username in the email itself; email accounts with usernames that were standard English terms were generally very highly priced.

In order to reduce the likelihood of identity theft, organizations that store personal information should take the necessary steps to protect data transmitted over the Internet or stored on their computers. This should include the development, implementation, and enforcement of secure policy requiring that all sensitive data is encrypted. Also, organizations should enforce compliance to information storage and transmission standards such as the PCI standard. This would ensure that even if the computer or medium on which the data were lost or stolen, the data would not be accessible. This step should be part of a broader security policy that organizations should develop and implement in order to ensure that any sensitive data is protected from unauthorized access.

Bot-infected computers

Bots are programs that are covertly installed on a user's machine in order to allow an unauthorized user to control the computer remotely. They allow an attacker to remotely control the targeted system through a communication channel such as IRC. These channels allow the remote attacker to control a large number of compromised computers in a bot network, which can then be used to launch coordinated attacks.

Bots allow for a wide range of functionality and most can be updated to assume new functionality by downloading new code and features. Bots can be used by external attackers to perform DoS attacks against an organization's Web site. Furthermore, bots within an organization's network can be used to attack other organizations' Web sites, which can have serious business and legal consequences. Bots can also be used by attackers to harvest confidential information from compromised computers, which can lead to identity theft. Furthermore, they can be used to distribute spam and phishing attacks, as well as spyware and adware.

An active bot-infected computer is one that carries out at least one attack per day. This does not have to be continuous; rather, a single computer can be active on a number of different days. Between January 1 and June 30, 2007, Symantec observed an average of 52,771 active bot-infected computers per day (figure 14), a 17 percent decrease from the previous reporting period.

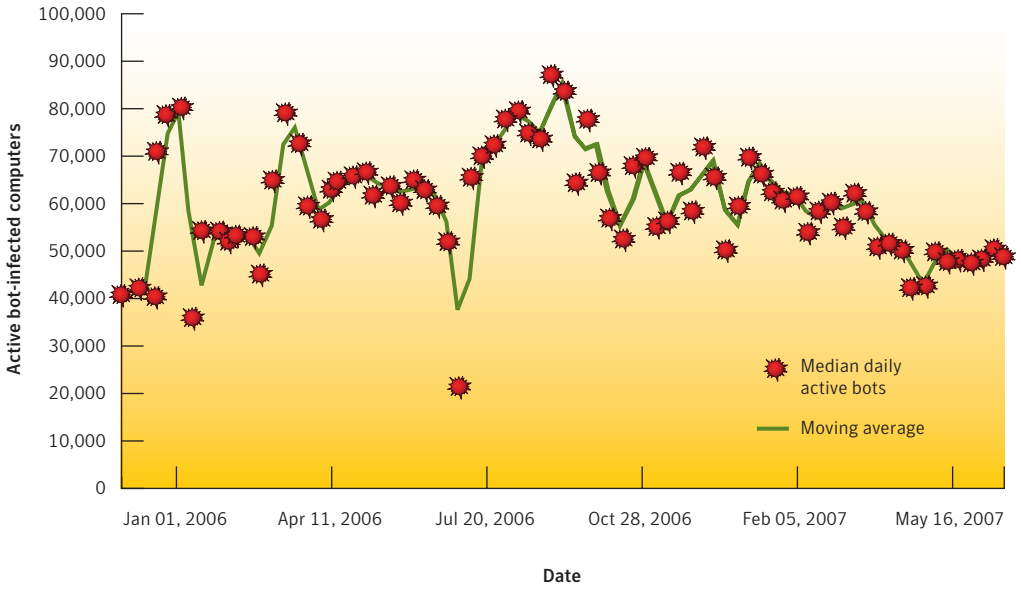

Figure 14. Active bot-infected computers per day
Source: Symantec Corporation

A distinct bot-infected computer is a distinct computer that was active at least once during the period. Symantec also observed 5,029,309 distinct bot-infected computers during this period, a 17 percent decrease from the last six months of 2006.

The decrease in bots observed over the past six months is likely due to a number of reasons, the primary one likely being a change in bot attack methods. As has been discussed in previous volumes of the Symantec *Internet Security Threat Report*, the exploitation of network-based vulnerabilities to spread bots is being slowly abandoned for methods that are more likely to succeed, such as bots that send a mass mailing of themselves.[85] Network-based attacks have been limited somewhat by the introduction of default firewalls in popular operating systems such as Microsoft Windows XP, as well as an increasing awareness of computer security issues among organizations and computer users. As a result, their use has declined, which has had the effect of limiting the propagation of bots.

Furthermore, law enforcement initiatives targeting bot-networks may also be having some effect. Recently the Federal Bureau of Investigation (FBI) in the United States released information on Operation Bot Roast. This is an ongoing cyber-crime initiative aimed at dismantling bot networks by identifying and arresting bot network owners and taking down the command-and-control servers by which they control their networks.[86] Initiatives such as these will likely result in a reduction in bots for a number of reasons. Firstly, as bot networks are dismantled, less bot activity will be observed. Secondly, as bot network owners become aware of the scrutiny of law enforcement agencies, they are likely to alter their tactics to avoid detection.

[85] For instance, please see Symantec *Internet Security Threat Report*, Volume IX (March 2006): http://eval.symantec.com/mktginfo/enterprise/white_papers/ent-whitepaper_symantec_internet_security_threat_report_ix.pdf : p. 30
[86] http://www.fbi.gov/pressrel/pressrel07/botnet061307.htm

Lifespan of bot-infected computers

For the first time, in this volume of the *Internet Security Threat Report*, Symantec is assessing the average lifespan of bot-infected computers. The lifespan of a bot is defined as the amount of time that elapses between the first detection of a bot-infected computer and the time that the computer is no longer actively attacking for 30 days, after which time it is assumed to have been disinfected. Gauging the average lifespan of bot-infected computers is important because it allows Symantec to assess how long bot-infected computers are present on a particular network prior to removal.

During the first six months of 2007, the lifespan of the average bot-infected computer was four days (figure 15). This is an increase from the previous period, when the average lifespan was three days. The median lifespan of a bot-infected computer during both periods was one day. This indicates that the majority of bot-infected computers only participate in attacking behavior for a short period, after which they are either identified and disinfected, or are used for activities other than carrying out Internet attacks, such as hosting spam zombies or phishing Web sites. The longest lifespan of a bot-infected computer during the period was 3.2 years. However, bots with such long life spans are rare.

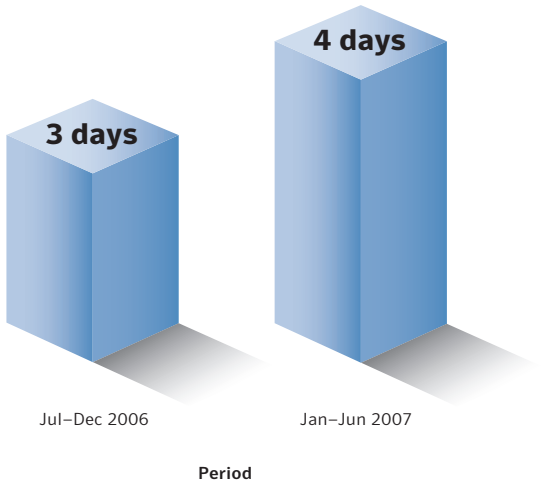

Figure 15. Average lifespan of bot-infected computers
Source: Symantec Corporation

The increase in the average lifespan from three to four days over the first six months of 2007 is not likely indicative of a fundamental change. Since the median remained the same, the change in overall average is driven by the longer-lasting bot-infected computers. The increased lifespan of the longer-lasting bot-infected computers has thereby increased the average lifespan. Thus, the bot lifespan is holding steady.

It therefore appears that law enforcement efforts, such as the FBI Bot Roast discussed above, as well as other security measures designed to identify and disinfect bot-infected computers are not reducing the lifespan of bot-infected computers. This is likely because the focus of those methods is to eliminate infections and keep infected computers free of bot software, and not necessarily to shorten the effective lives of bot-infected computers. This is supported by the fact that the number of bot-infected computers has decreased during the period while their lifespan remains steady.

Symantec also tracks the number of bot command-and-control servers worldwide. Command-and-control servers are computers that bot network owners use to relay commands to bot-infected computers on their networks, usually through IRC channels. In the first six months of 2007, Symantec identified 4,622 bot command-and-control servers (figure 16). This is a three percent decrease from the previous period, when 4,746 command-and-control servers were identified.

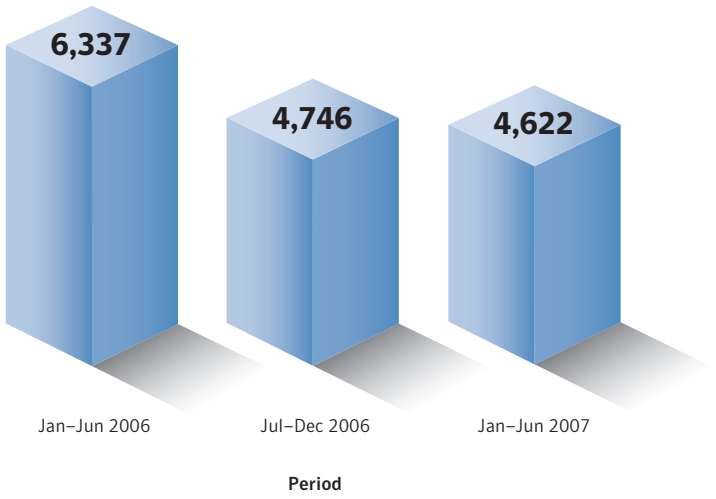

Figure 16. Command-and-control servers
Source: Symantec Corporation

The decrease in command-and-control servers reflects a consolidation of bot networks that Symantec first observed in second half of 2006.[87] During that period, the number of command-and-control servers decreased and the average size of bot networks increased. As a result, over the past year, bot networks appear to have become more concentrated in the hands of fewer bot network owners.

[87] Symantec *Internet Security Threat Report*, Volume XI (March 2007): http://eval.symantec.com/mktginfo/enterprise/white_papers/ent-whitepaper_internet_security_threat_report_xi_03_2007.en-us.pdf : p. 17, 34

The marginal drop observed in the first six months of 2007 is likely due to a change in the fundamental methods that bots use to communicate. That is, bot network owners are moving away from using command-and-control servers and adopting new methods instead. One example is the fast flux domain name service scheme.[88] In this scheme, control of bot networks is diffused through a number of computers within the bot network. This removed the need for a single command-and-control server, and as such may represent a future trend that will make command-and-control servers less common. Other trends in methods of communication such as peer-to-peer communication will also lend to the decrease in the number of command-and-control servers.

To prevent bot infections, Symantec recommends that ISPs perform both ingress and egress filtering to block known bot traffic. ISPs should also filter out potentially malicious email attachments to reduce exposure to enterprises and end users.

Organizations should monitor all network-connected computers for signs of bot infection, ensuring that any infections are detected and isolated as soon as possible. They should also ensure that all antivirus definitions are updated regularly. As compromised computers can be a threat to other systems, Symantec also recommends that enterprises notify their ISPs of any potentially malicious activity.

Organizations should also perform egress filtering on outgoing network traffic, ensuring that malicious activity and unauthorized communications are not taking place. They should also create and enforce policies that identify and restrict applications that can access the network.

To reduce exposure to bot-related attacks, end users should employ defense-in-depth strategies, including the deployment of antivirus software and a firewall. Users should update antivirus definitions regularly and ensure that all desktop, laptop, and server computers are updated with all necessary security patches from their operating system vendor. Symantec also advises that users never view, open, or execute any email attachment unless the attachment is expected and comes from a known and trusted source, and unless the purpose of the attachment is known.

Bot-infected computers by country

Recognizing the ongoing threat posed by bot networks, Symantec tracks the distribution of bot-infected computers worldwide. This can help analysts understand how bot-infected computers, and the networks they constitute, are distributed globally. This is important, as a high percentage of bot-infected computers likely indicates a greater potential for bot-related attacks. It could also give insight into the level of patching and security awareness amongst computer administrators and users in a given region, as initial bot infections usually take advantage of unpatched computer systems.

China had the highest number of bot-infected computers during the first half of 2007, accounting for 29 percent of the worldwide total (figure 17). This is a slight increase from 26 percent in the second half of 2006, when China also had the highest number of bot-infected computers.

[88] http://www.securityfocus.com/news/11473

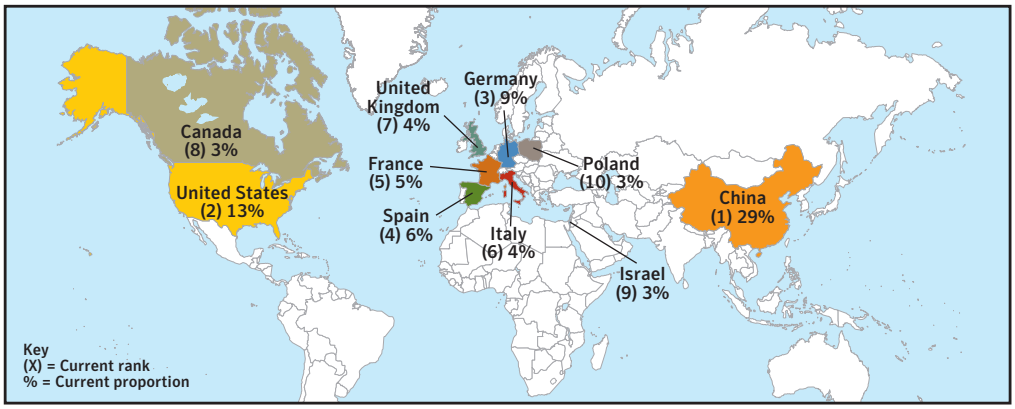

Figure 17. Bot-infected computers by country
Source: Symantec Corporation

Symantec has observed that bots usually infect computers that are connected to high-speed broadband Internet through large ISPs and that the expansion of broadband connectivity often facilitates the spread of bots. China's Internet infrastructure is currently expanding rapidly.[89] Between May 2006 and May 2007, China added more broadband lines than any other country.[90]

However, it is worth noting that China's increase in bot-infected computers seems to be slowing. In the first half of 2006, the percentage of worldwide bot-infected computers situated in China increased from nine percent to 20 percent. In the second half of 2006, the rate of increase slowed to six percentage points, from 20 percent to 26 percent. In the first half of 2007, it went up only three percentage points. This may be a sign that security awareness, practices and infrastructure are beginning to catch up with the rapid growth of Internet usage in China.

In the first six months of 2007, the United States had the second highest number of bot-infected computers, accounting for 13 percent of the worldwide total. This is almost unchanged from the second half of 2006, when the United States ranked second, accounting for 14 percent of the world's bot-infected computers. Germany had the third highest number of bot-infected computers during the first half of 2007, accounting for nine percent of the worldwide total. During the second half of 2006, Germany ranked fourth and accounted for six percent of the world's bot-infected computers.

During the first half of 2007, the United States had the most known command-and-control servers worldwide (table 4), accounting for 43 percent of the worldwide total. This is a marginal increase from the previous period, when the United States was also the site of the most command-and-control servers, accounting for 40 percent of the worldwide total.

[89] http://www.vnunet.com/vnunet/news/2163552/china-lead-broadband-world
[90] http://www.point-topic.com

The high proportion of command-and-control servers in the United States likely indicates that servers there control not only bot networks within the country but elsewhere as well. The high proportion of bot-infected computers and command-and-control servers in the United States is driven by that country's extensive Internet and technology infrastructure. As of June 2006, more than 58 million broadband Internet users were located there, the highest number in the world.[91]

Current Rank	Previous Rank	Country	Current Proportion	Previous Proportion
1	1	United States	43%	40%
2	3	Germany	7%	6%
3	5	Canada	7%	4%
4	2	South Korea	6%	10%
5	4	China	3%	5%
6	9	United Kingdom	3%	2%
7	6	Taiwan	3%	3%
8	10	Italy	2%	2%
9	7	Sweden	2%	3%
10	11	Turkey	2%	2%

Table 4. Command-and-control servers by country
Source: Symantec Corporation

Germany had the second highest number of command-and-control servers in the first six months of 2007, accounting for seven percent of the worldwide total. During the previous period, Germany ranked third and accounted for six percent of the worldwide total. During the current reporting period, Canada had the third most command-and-control servers in the world, accounting for seven percent of the total. This is an increase over the second half of 2006, when Canada ranked fifth and accounted for four percent of the world's total.

[91] http://www.oecd.org/document/7/0,3343,en_2649_34223_38446855_1_1_1_1,00.html

Vulnerability Trends

Vulnerabilities are design or implementation errors in information systems that can result in a compromise of the confidentiality, integrity, or availability of information stored upon or transmitted over the affected system. They are most often found in software; however, they exist in all layers of information systems, from design or protocol specifications to physical hardware implementations. Vulnerabilities may be triggered actively—either by malicious users or automated malicious code—or passively during system operation. The discovery and disclosure of a single vulnerability in a critical asset can seriously undermine the security posture of an organization.

New vulnerabilities are discovered and disclosed regularly by a sizeable community of end users, security researchers, hackers, security vendors, and occasionally by the software vendors themselves. Symantec carefully monitors vulnerability research, tracking vulnerabilities throughout their lifecycle, from initial disclosure and discussion to the development and release of a patch or other remediation measure.

Symantec operates one of the most popular forums for the disclosure and discussion of vulnerabilities on the Internet, the BugTraq mailing list, which has approximately 50,000 direct subscribers who contribute, receive, and discuss vulnerability research on a daily basis.[92] Symantec also maintains one of the world's most comprehensive vulnerability databases, currently consisting of over 22,000 vulnerabilities (spanning more than a decade) affecting more than 50,000 technologies from over 8,000 vendors. The following discussion of vulnerability trends is based on a thorough analysis of that data.

This section of the Symantec *Internet Security Threat Report* will discuss vulnerabilities that have been disclosed between January 1 and June 30, 2007. It will compare them with those disclosed in the previous six-month period and discuss how current vulnerability trends may affect potential future Internet security activity.

Vulnerability trends highlights

The following section will offer a brief summary of some of the vulnerability trends that Symantec observed during this reporting period based on data provided by the sources listed above. Following this overview, the *Internet Security Threat Report* will discuss selected vulnerability metrics in greater depth, providing analysis and discussion of the trends indicated by the data.

- Symantec documented 2,461 vulnerabilities in the first half of 2007, three percent less than the second half of 2006.
- Symantec classified nine percent of all vulnerabilities disclosed during this period as high severity, 51 percent were medium severity, and 40 percent were low. In the second half of 2006, four percent of newly disclosed vulnerabilities were high severity, 69 percent were medium severity, and 27 percent were low severity.
- Sixty-one percent of vulnerabilities disclosed during this period affected Web applications, down from 66 percent in the second half of 2006.

[92] The BugTraq mailing list is hosted by SecurityFocus (http://www.securityfocus.com). Archives are available at http://www.securityfocus.com/archive/1

- Seventy-two percent of vulnerabilities documented in this reporting period were easily exploitable. This is a decrease from 79 percent in the previous reporting period.
- In the first half of 2007, all operating systems except Hewlett Packard HP-UX had shorter average patch development times than in the second half of 2006.
- Hewlett-Packard HP-UX had an average patch development time of 112 days in the first half of 2007, the highest of any operating system. Sun had the highest average patch development time in the second half of 2006, with 145 days.
- The average window of exposure for vulnerabilities affecting enterprise vendors was 55 days. This is an increase over the 47-day average in the second half of 2006.
- Symantec documented 39 vulnerabilities in Microsoft Internet Explorer, 34 in Mozilla browsers, 25 in Apple Safari, and seven in Opera. In the second half of 2006, 54 vulnerabilities were disclosed for Internet Explorer, 40 for Mozilla browsers, four for Apple Safari, and four for Opera.
- Apple Safari had an average window of exposure of three days in the first half of 2007, the shortest of any browser reviewed during this period. Mozilla browsers had the shortest average window of exposure in the second half of 2006, two days.
- Symantec documented six zero-day vulnerabilities in the first half of 2007, down from the 12 that were reported during the second half of 2006.
- Ninety-seven vulnerabilities were documented in Oracle, more than any other database during the first half of 2007. Oracle also had the most database vulnerabilities in the second half of 2006, with 168.
- There were 90 unpatched enterprise vendor vulnerabilities in the first half of 2007, which is down from the 94 documented in the second half of 2006. Microsoft had the most unpatched vulnerabilities of any enterprise vendor during both of these periods.
- In the first half of 2007, Symantec documented 237 vulnerabilities in Web browser plug-ins. This is a significant increase over 74 in the second half of 2006.
- During the first half of 2007, 89 percent of plug-in vulnerabilities disclosed affected ActiveX components for Internet Explorer. ActiveX components accounted for 58 percent of plug-in vulnerabilities in the second half of 2006.
- Symantec found that more than 50 percent of medium- and high-severity vulnerabilities patched by operating system vendors affected Web browsers or had other client-side attack vectors during this and the previous reporting period. Apple was the sole exception, with 49 percent of the vulnerabilities examined in the first half of 2007 affecting browsers or having client-side attack vectors.

Vulnerability Trends Discussion

This section will discuss selected vulnerability trends in greater depth, providing analysis and discussion of the trends indicated by the data. The following metrics will be discussed:

- Patch development time for operating systems
- Patched operating system vulnerability by type
- Window of exposure for enterprise vendors
- Web application vulnerabilities
- Web browser vulnerabilities
- Window of exposure for Web browsers
- Zero-day vulnerabilities
- Unpatched enterprise vendor vulnerabilities
- Browser plug-in vulnerabilities
- Vulnerabilities—protection and mitigation

Patch development time for operating systems

The time period between the disclosure date of a vulnerability and the release date of an associated patch is known as the patch development time. If exploit code is created and made public during this time, computers may be immediately vulnerable to widespread attack. This metric will assess and compare the average patch development times of medium- and high-severity vulnerabilities affecting five different operating systems: Apple Mac OS® X, Hewlett-Packard HP-UX, Microsoft Windows, Red Hat® Linux (including enterprise versions and Red Hat Fedora), and Sun Microsystems Solaris™.

Of the five operating systems tracked in the first six months of 2007 (figure 18), Microsoft had the shortest average patch development time at 18 days, based on a sample set of 38 patched vulnerabilities. Of the 38 vulnerabilities, two affected third-party applications. This is lower than the average patch development time of 23 days in the second half of 2006 based on a sample set of 50 vulnerabilities, seven of which affected third-party applications.

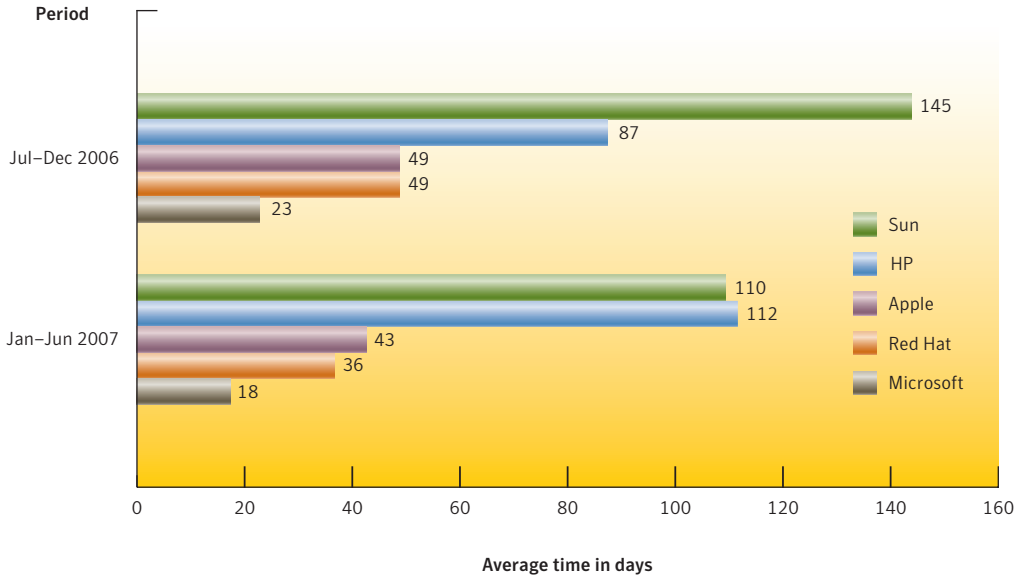

Figure 18. Patch development time for operating systems
Source: Symantec Corporation

Red Hat had the second shortest average patch development time in the first six months of 2007, with an average of 36 days for a sample set of 91 vulnerabilities. Of these, 90 affected third-party applications. The average patch development time is down from 49 days in the second half of 2006, which was based on 149 vulnerabilities, all of which affected third-party applications.

Apple had the third shortest average patch development time in the first half of 2007; it was 43 days for a sample set of 59 vulnerabilities. Nine of those vulnerabilities affected third-party applications. This is a shorter average patch development time than the 49 days reported in the second half of 2006, which was based on a sample set of 32 vulnerabilities, including 12 that affected third-party applications.

Sun had the fourth shortest average patch development time in the first half of 2007, at 110 days for a sample set of 73 vulnerabilities. Sixty-seven of those affected third-party applications. This figure is down from the 145 day patch development time in the second half of 2006. This was based on a sample set of 35 vulnerabilities, 32 of which affected third-party applications.

HP had the longest average patch development time during this reporting period, at 112 days. This was based on a total of 30 vulnerabilities, 28 of which affected third-party applications. The average patch development time for this period was higher than the 87 days reported in the second half of 2006. The previous period was based on a sample set of 70 vulnerabilities, 68 of which affected third-party applications.

Vulnerabilities affecting third-party applications are still a factor in the average patch development time for operating systems. Vendors with fewer third-party applications to patch generally have an advantage over those whose operating systems comprise many third-party components. However, the vulnerabilities affecting these vendors often affect core proprietary components; therefore, the operating systems are more likely to be vulnerable in their default installation.

The numbers from this and previous volumes of the report demonstrate that Red Hat has had the best track record in dealing with third-party vulnerabilities. This may be due to the extent of their involvement with third-party vendors and the open-source community, as they often contribute their own patches and work closely with third-party vendors.

Patched operating system vulnerability by type

In this version of the *Internet Security Threat Report*, Symantec will be discussing the types of vulnerabilities that are assessed in the "Patch development time for operating systems" metric. It will also consider vulnerabilities affecting the same group of vendors. This will provide insight into the types of applications and vulnerabilities that are present in the operating systems that are examined in the previous metric.

The sample sets are limited to vulnerabilities that are considered medium or high severity. Vulnerabilities are divided into the following categories:

- Web browser
- Client-side
- Local
- Server

Some vulnerabilities did not fit into these categories and these cases are noted in the discussion.

Of the 59 patched vulnerabilities that affected Apple Mac OS X in the first half of 2007, eight affected browsers, 21 were client-side vulnerabilities, 17 were local, 11 affected servers, and two vulnerabilities did not fit into any of these categories (figure 19). During the last six months of 2006, Apple had one patched browser vulnerability, 18 client-side vulnerabilities, seven that were local, four in servers, and two that could not be categorized.

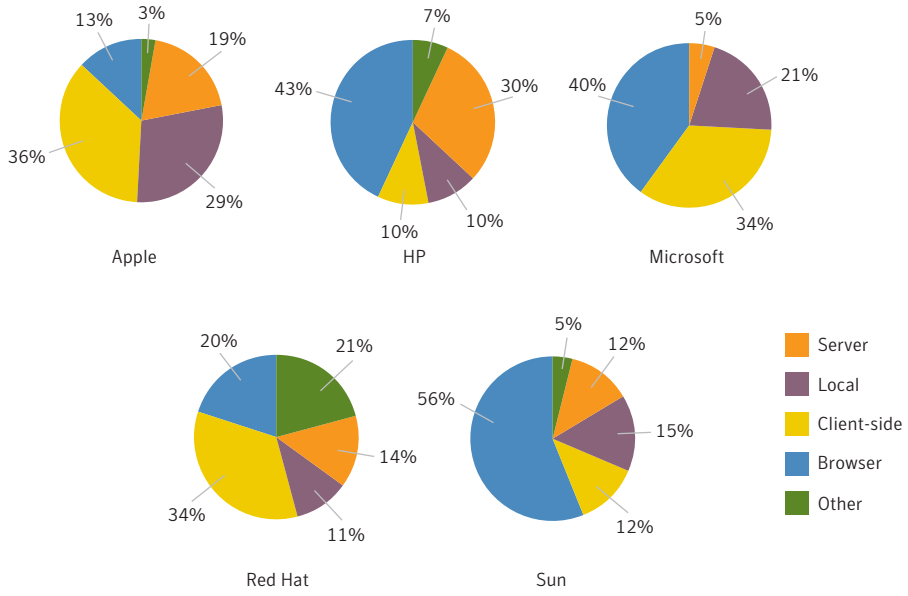

Figure 19. Patched operating system vulnerability by type
Source: Symantec Corporation

There were 30 patched vulnerabilities disclosed during this period that affected HP-UX. Of these, 13 affected browsers, three were client-side, three were local, nine affected servers, and two could not be categorized. From a sample set of 70 patched vulnerabilities in the second half of 2006, 50 affected browsers, four were client-side issues, one was locally exploitable, 13 affected servers, and two fell outside of these categories.

In the first half of 2007, Symantec disclosed 38 vulnerabilities for Microsoft Windows that were patched. Fifteen affected Web browsers, 13 were client-side issues, eight were locally exploitable, and two affected servers. The 50 vulnerabilities patched by Microsoft during the second half of 2006 consisted of 15 browser issues, 20 client-side vulnerabilities, three issues that were local, and 12 that affected servers.

The set of patched vulnerabilities for Red Hat Linux during this reporting period consisted of 91 vulnerabilities. Eighteen of these issues affected browsers, 31 were client-side, 10 were local, and 13 affected servers. The remaining 19 were unclassifiable according to the criteria for this metric. Of the 149 Red Hat Linux vulnerabilities in the previous reporting period, 47 affected browsers, 53 were client-side issues, 22 were local, 12 affected servers, and 15 did not fit into any of these categories.

Of 73 patched vulnerabilities in Sun Solaris during the first six months of 2007, 41 affected browsers, nine were client-side issues, 11 were local, nine affected servers, and three could not be categorized. During the second half of 2006, 35 patched vulnerabilities were categorized. Of these, 25 affected browsers, one was a client-side vulnerability, four were local, and four affected servers. One vulnerability could not be categorized.

For all vendors, the majority of patched vulnerabilities affected Web browsers or were client-side issues. Browser and client-side vulnerabilities are similar in that they typically require a user to interact with malicious content, whether it is a Web-page or a malicious file. As such, the attacker must usually present the content to the user in a manner that is enticing and seems innocuous. This tactic is typical of targeted attacks, which may be directed at users within a specific organization or who visit a particular Web site that the attacker has compromised.

Exploits of browser and client-side vulnerabilities may not necessarily result in a complete compromise of the affected computer. This is because they can only perform actions in the context of the currently logged-in user, who may not possess administrative access. In previous issues of this report, Symantec has emphasized a shift from attacks that target servers or network assets to those that target desktop users through a myriad of application-level vulnerabilities. The data for this metric demonstrates that these types of applications appear to be a priority for security researchers and attackers. Since all of the vulnerabilities that were examined are patched, it also shows that vendors are responding to this threat.

Window of exposure for enterprise vendors

Attackers use custom-developed code known as exploit code, or exploits, to take advantage of vulnerabilities to compromise a computer. The time lapse between the publication of an initial vulnerability report and the appearance of third-party exploit code is known as the "exploit code development time."[93] This is a concern to enterprises because it is a measurement of how long it takes for the average exploit to become public. If an exploit is published before a patch is available, administrators must implement other protective measures to reduce the risk of attack.

It is important to note that the set of vulnerabilities included in this metric is limited and does not represent all software from all possible vendors. Instead, it only includes vendors that are classified as enterprise vendors. The purpose is to illustrate the window of exposure for widely deployed, mission-critical software. Because of the large number of vendors with technologies that have a very low deployment, only exploits for technologies from enterprise vendors are included.[94]

In the first half of 2007, the window of exposure for enterprise vendors was 55 days. This was based on an average exploit development time of six days and an average patch development time of 61 days. The enterprise window of exposure for the second half of 2006 was 47 days. The average exploit development time was five days and the average patch development time for enterprise vendors was 52 days.

The window of exposure has risen over the last three reporting periods.[95] This is primarily due to the influence of longer patch times required for vulnerabilities that affect third-party components in some operating system vendors, such as browser plug-ins (which are discussed in the "Browser plug-in vulnerabilities" metric below). Compared to operating system vendors, other vendors have a relatively short average patch development time. Non-operating system vendors are less dependent on the developers of third-party components to develop patches for vulnerabilities in their products. This gives them an advantage over vendors who distribute and maintain products containing third-party applications.

[93] It should be noted that the data included in this discussion is limited to public examples of exploit code that Symantec has associated with specific vulnerabilities. There are many instances in which a private or commercial exploit may be available, but this data cannot be consistently tracked since exploit publication dates are not available.
[94] Vendors included in this metric are: Microsoft, Sun, HP, Symantec, EMC, IBM, Cisco, Oracle, CA (Computer Associates), and McAfee.
[95] For a discussion of the window of exposure for enterprise vendors in the first half of 2006, please see the Symantec *Internet Security Threat Report*, Volume X (September 2006): http://eval.symantec.com/mktginfo/enterprise/white_papers/ent-whitepaper_symantec_internet_security_threat_report_x_09_2006.en-us.pdf : p. 58

Other enterprise vendors, including security vendors, have demonstrated better responsiveness to vulnerabilities than the operating system vendors. Vendor responsiveness is especially important to security vendors, who are often targeted by security researchers and attackers in order to either improve the security products or damage the credibility of such vendors.

Web application vulnerabilities

Web applications are technologies that use a browser for their user interface, rely on HTTP as the transport protocol, and reside on Web servers. Examples of Web-based applications include content management systems, e-commerce suites (such as shopping cart implementations), Weblogs, and Web-based email.

The online presence of an organization is often facilitated through Web applications, particularly as an increasing number of traditional software vendors are bolstering their existing applications with Web-based user interfaces, or converting them over entirely. Web applications may be the site of vulnerabilities that can be exploited to gain unauthorized access to computers on which they are deployed. Users within the organization may also be affected by insecure Web sites, which may present a risk of compromise and/or a threat to confidential information.

In the first half of 2007, 61 percent of all vulnerabilities affected Web applications (figure 20). This is a drop from the 66 percent reported in the second half of 2006, and a further decrease from the 69 percent of all vulnerabilities that affected Web applications in the first half of 2006.

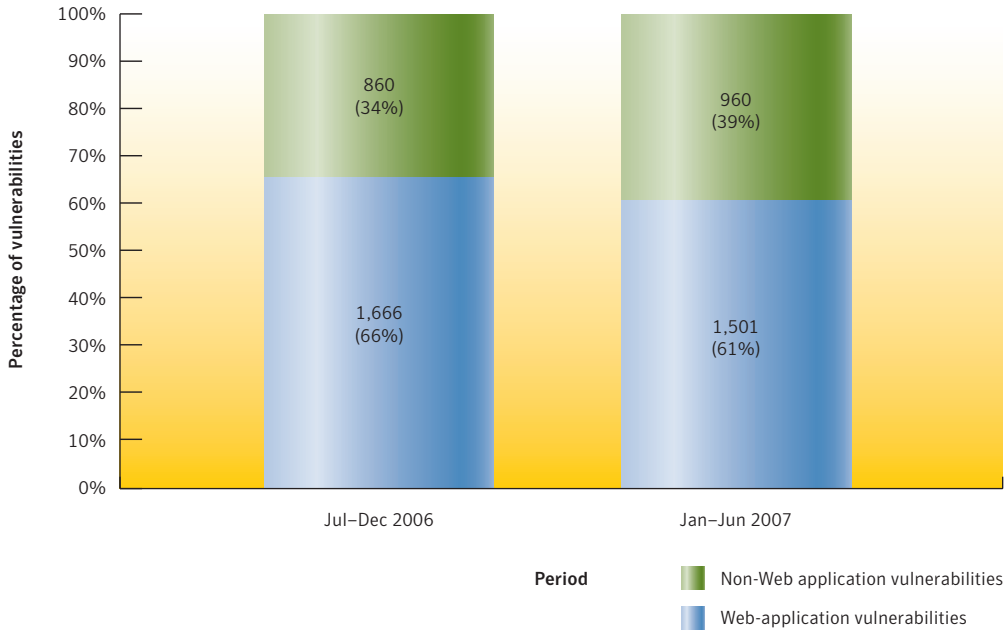

Figure 20. Web application vulnerabilities
Source: Symantec Corporation

Prior to this volume of the *Internet Security Threat Report*, Symantec had observed that the proportion of Web application vulnerabilities had been on the rise. This trend persisted for five reporting periods starting in the first half of 2004 and ending in the first half of 2006. Because of their increasing prevalence, Web application vulnerabilities appear to have influenced other vulnerability trends discussed in this report during each period.

This is true for the current period as well. For instance, the decrease in Web application vulnerabilities during the first six months of 2007 has contributed to the drop in the total number of vulnerabilities documented this period. As well, Web application vulnerabilities are typically classified as easily exploitable; therefore, the current decrease likely accounts for the drop in easily exploitable vulnerabilities over the past six months. Furthermore, a lower percentage of Web application vulnerabilities may have resulted in a higher percentage of high-severity vulnerabilities. (Each of these metrics will be discussed subsequently in this section.)

The decrease in Web application vulnerabilities that was observed during this period may be due to security researchers focusing more of their efforts on finding vulnerabilities that are specific to a particular Web site. These site-specific vulnerabilities are often discovered during an unauthorized audit of the Web site and usually require the same amount of research effort as other Web application vulnerabilities.

The legality of discovering and disclosing site-specific vulnerabilities is in question, as well, because it often requires that the researcher performs attacks on the affected site.[96] Both security researchers and attackers have various incentives for seeking out site-specific vulnerabilities. Researchers may garner more attention for themselves if they report a vulnerability in a popular Web site than if they discover a similar vulnerability in a lesser-known Web application. These vulnerabilities are also appealing to attackers because they may provide a means of compromising a Web site that can be employed in other attacks. In such a scenario, the attacker may use the legitimacy of the Web site to attract victims of subsequent attacks. Sites with large user bases, such as MySpace, have already been abused in this manner.[97]

In the first half of 2007, security researchers staged a "Month of MySpace Bugs"[98] and a "Month of Search Engine Bugs"[99] to bring various site-specific vulnerabilities into the public spotlight. However, because of the legal concerns for the researchers reporting these issues and for any site or database that collects reports of these issues, it may be difficult to verify the number of legitimate site-specific vulnerabilities that are being discovered and reported. In addition, when the administrator of a Web site patches a site-specific vulnerability, it no longer exists. As such, Symantec has no insight into the number of site-specific vulnerabilities that are being discovered and reported. However, some resources have emerged to facilitate the full disclosure of site-specific vulnerabilities.[100] There have also been public incidents that suggest that attackers are discovering these vulnerabilities in bulk.[101]

Web application vulnerabilities are also likely candidates for multistaged attacks. During the first half of 2007, an unspecified cPanel exploit was used to compromise legitimate Web sites hosted through a common Web hosting provider. These were then used by MPack[102] to launch client-side exploits on unsuspecting users.[103] It is also possible that attackers could exploit this or a similar vulnerability to set up phishing Web sites or other malicious sites through the Web-hosting provider. Vulnerabilities that let attackers inject arbitrary content such as cross-site scripting into Web sites may be employed in a similar manner to launch attacks against users of legitimate sites.

[96] http://www.darkreading.com/document.asp?doc_id=125984&WT.svl=news1_1
[97] http://blog.washingtonpost.com/securityfix/2007/06/web_2pointuhoh_worm_whacks_mys.html
[98] http://momby.livejournal.com/7285.html
[99] http://Websecurity.com.ua/1114
[100] http://www.xssed.com
[101] http://www.theregister.co.uk/2007/06/20/youtube_security_ultimatum
[102] MPack is a commercially available black market attack toolkit. It can launch exploits for browser and client-side vulnerabilities against users who visit a malicious or compromised Web site. For more information, see http://www.symantec.com/enterprise/security_response/weblog/2007/05/mpack_packed_full_of_badness.html
[103] http://isc.sans.org/diary.html?storyid=3015&rss

Web browser vulnerabilities

The Web browser is a critical and ubiquitous application that has become an increasingly popular subject for vulnerability researchers over the past few years. Traditionally, the focus of security researchers has been on the perimeter: servers, firewalls, and other assets with external exposure. However, security researchers and attackers now consider client-side vulnerabilities to be a fruitful area of research and attacks. As part of this shift toward client-side issues, vulnerabilities in Web browsers have become increasingly prominent, which in turn pose a threat to end users' desktop computers.

Browsers are complex and feature rich, traits that can expose them to vulnerabilities in newly implemented features. Due to the integration of various content-handling applications—such as productivity suites and media players—browsers have become a viable attack vector for many client-side vulnerabilities. This is particularly true of operating systems in which the browser is not disassociated from many other operating system processes and features.

Web browser vulnerabilities are a serious security concern due to their role in online fraud and the propagation of spyware and adware. They are particularly prone to security concerns because they come in contact with more potentially untrusted or hostile content than other applications. This metric will examine vulnerabilities that were disclosed for a number of Web browsers during the first six months of 2007.

During this period, Symantec documented 39 vulnerabilities in Microsoft Internet Explorer (figure 21). Of these, one was considered to be high severity, 15 were medium severity, and 23 were low. This total is a decrease from the 54 vulnerabilities documented in the second half of 2006. Of those, one was considered high severity, 13 were medium severity, and 40 were low.

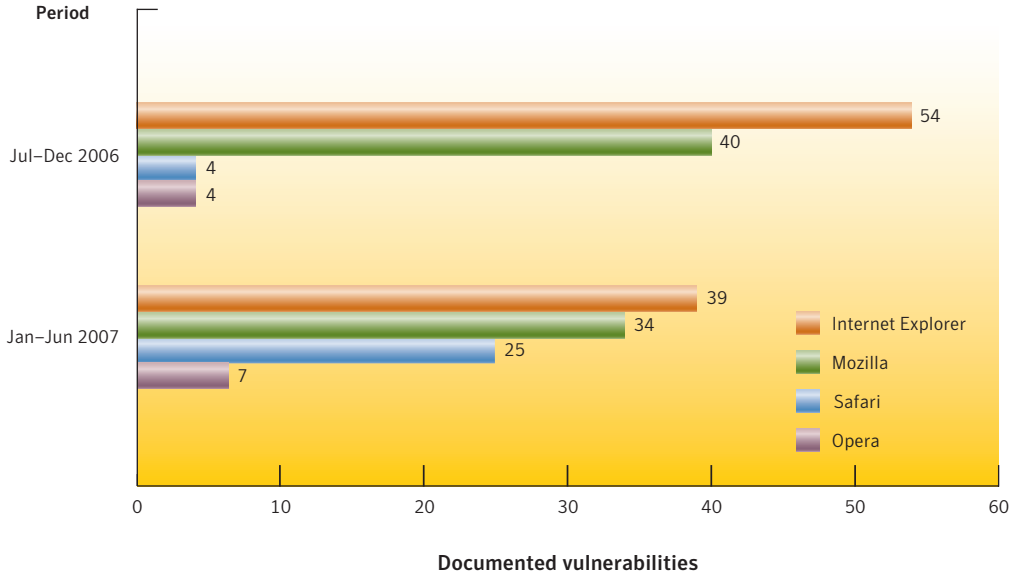

Figure 21. Web browser vulnerabilities
Source: Symantec Corporation

During the first half of 2007, 34 vulnerabilities were disclosed that affected Mozilla browsers. Of these, 12 were considered to be medium severity and 22 were considered low. This total is a decrease from the 40 vulnerabilities that affected Mozilla browsers in the second half of 2006. Of those, 35 were considered medium severity and five were low severity.

Safari was affected by 25 vulnerabilities in the first half of 2007. Seven of these were medium-severity vulnerabilities, and the other 18 were low severity. This is an increase from the four Safari vulnerabilities that were documented in the second half of 2006. Of these, two were medium severity and two were low severity.

In the first six months of 2007, Symantec documented seven vulnerabilities that affected Opera. Of these, three were medium severity and the other four were low. The total of seven is an increase from the four vulnerabilities that affected Opera in the second half of 2006, two of which were considered medium severity and two of which were low.

During the current reporting period, the majority of vulnerabilities documented in all browsers were low severity. These vulnerabilities consisted of denial of service, information disclosure, and spoofing issues. This may be indicative of improvements in the security of the current generation of browsers. It is possible that many of the higher-severity vulnerabilities have been discovered by the current generation of fuzzers.[104] Such improvements are likely to be short-lived due to the evolution of fuzzing techniques and competition among browser vendors to include more features that will likely expose new vulnerabilities. In spite of the trend towards lower-severity vulnerabilities, Web browsers are still implicated in attacks through vulnerabilities in browser plug-in attacks and client-side issues.

With the exception of denial of service vulnerabilities, many of the low-severity issues are still a concern, as they may facilitate phishing attacks or allow attackers to gain access to sensitive information. These vulnerabilities are symbolic because they represent subtle attacks against the security model of the browser. The current generation of browsers includes security features that are intended to protect users against attacks such as phishing. It is reasonable to speculate that these new security features may become the focus of security researchers and attackers alike.

Safari was subject to the greatest change in the number of vulnerabilities over previous reporting periods. During first half of 2007, Apple released beta versions of Safari for Windows.[105] This event drew the attention of security researchers, who discovered a number of vulnerabilities shortly after the release.[106] Beta software does not carry the same guarantees of security as production versions, so it is not surprising to see that vulnerabilities were quickly discovered in the beta Safari for Windows. However, some of the issues were also found to affect production Safari releases for Mac OS X. As Safari becomes more accessible and its market share increases, it is likely to receive more attention from security researchers and attackers.

[104] Fuzzing is a security research and quality assurance method that generally entails providing randomly generated inputs in an attempt to discover vulnerabilities and bugs. Fuzzers are programs or scripts that are designed to find vulnerabilities in software code or scripts. They have automated many of the code auditing tasks that security researchers had previously done manually.
[105] http://www.symantec.com/enterprise/security_response/weblog/2007/06/new_technologies_from_apple.html
[106] http://www.symantec.com/enterprise/security_response/weblog/2007/06/vulnerabilities_for_safari_on.html

In order to protect against successful exploitation of Web browser vulnerabilities, Symantec advises administrators and end users to upgrade all browsers to the latest, patched versions. Symantec recommends that organizations educate users to be extremely cautious about visiting unknown or untrusted Web sites and viewing or following links in unsolicited emails. Administrators should also deploy Web proxies in order to block potentially malicious script code. Administrators and end users should actively maintain a white-list of trusted sites and disable individual plug-ins and scripting capabilities for all other sites. This will not prevent exploitation attempts from white-listed sites, but may aid in preventing exploits from all other sites. Organizations can also implement a white-list policy at the network perimeter to regulate outgoing access by end users.

Window of exposure for Web browsers

The window of exposure is the difference in days between the time at which exploit code affecting a vulnerability is made public and the time at which the affected vendor makes a patch available to the public for that vulnerability. During this time, the computer or system on which the affected application is deployed may be susceptible to attack, as administrators will have no official recourse against exploitation of the vulnerability. Instead they will have to resort to best practices and workarounds to reduce the risk of successful compromise.

This metric will assess the window of exposure for vulnerabilities in selected Web browsers. For this version of the *Internet Security Threat Report*, Symantec will be supplementing the Web browser window of exposure discussion with the maximum amount of time that elapsed between the disclosure of a single vulnerability and the release of an associated patch. Maximum patch times indicate the longest period of time required for a patch to be released to the public.

During the first half of 2007, Apple Safari had a window of exposure of three days, a decrease over the 62-day window in the second half of 2006 (figure 22). The window of exposure for the first half of 2007 was based on a sample set of 13 vulnerabilities, with a maximum patch time of eight days. The results for the second half of 2006 were based on a sample set of one vulnerability with a patch time of 62 days.

Safari had the smallest window of exposure of any browser. As discussed in the “Browser vulnerabilities” section of this report, Apple released the Safari for Windows beta. A number of vulnerabilities were discovered in the browser shortly after its release. The quick response to these vulnerabilities by Apple resulted in a shorter window of exposure.

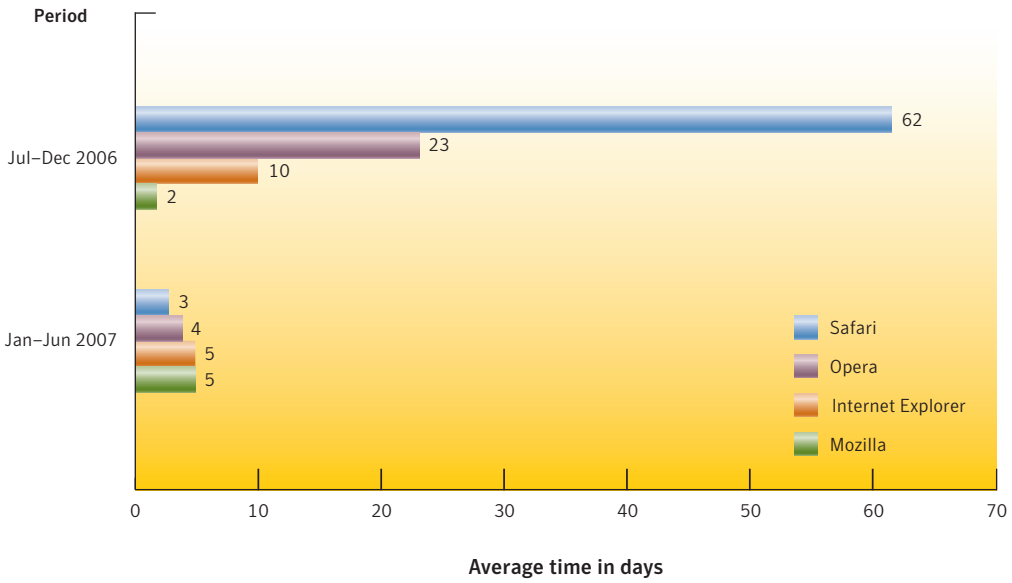

Figure 22. Window of exposure for Web browsers
Source: Symantec Corporation

In the first six months of 2007, Opera had a window of exposure of four days based on a sample set of five patched vulnerabilities. This is an increase over the 23-day window in the second half of 2006, which was based on a sample set of three patched vulnerabilities. In the current reporting period, Opera had maximum patch development time of 23 days. This can be attributed to a few vulnerabilities in a small sample data set that disproportionately affected the average. In the previous six-month period, a maximum of 46 days elapsed before a patch was available for vulnerabilities in Opera.

In the first half of 2007, Microsoft Internet Explorer had a window of exposure of five days based on a sample set of 17 patched vulnerabilities. This is a decrease from the 10-day time period in the second half of 2006, which was based on a sample set of 15 patched vulnerabilities. The maximum patch development time for Internet Explorer vulnerabilities during the current reporting period was 90 days. In the second half of 2006, the maximum patch development time was 78 days.

During the first six months of 2007, Mozilla had a window of exposure of five days based on a sample set of 22 patched vulnerabilities. This is an increase over the window of exposure of two days in the second half of 2006, which was based on 36 patched vulnerabilities. During the current reporting period, Mozilla had a maximum patch development time of 83 days. In the second half of the year, the maximum patch development time was 33 days.

With the exception of Mozilla, all the Web browser vendors had a shorter window of exposure in the first half of 2007. However, readers should note that Opera and Safari figures for the last six months of 2006 were skewed by small sample sets, so this may a factor in their shorter window of exposure during the current period.

Exploitation of Internet Explorer and ActiveX vulnerabilities in the wild may have contributed to the shorter window of exposure for Internet Explorer. The majority of Internet Explorer vulnerabilities in this period were announced by the vendor and patched when they were announced. Slower patch times for lower-severity vulnerabilities were a factor in the longer window of exposure for Mozilla.

The average time for an exploit to emerge remains minimal; many exploits are released at the same time that vulnerabilities are announced or shortly afterwards. The low average patch development times indicate that vendors are also quick to respond, which is likely due to the high likelihood that browser vulnerabilities will be exploited in the wild. Responsible disclosure efforts contribute to these low numbers, as many of the vulnerabilities were discovered by third-parties but publicly announced only when patches were made available.

Zero-day vulnerabilities

A zero-day vulnerability is one that appears to have been exploited in the wild prior to being publicly known. It may not have been known to the vendor prior to exploitation, and the vendor had not released a patch at the time of the exploit activity.

Zero-day vulnerabilities represent a serious threat in many cases, because there is no patch available for them and because they will likely be able to evade purely signature-based detection. It is the unexpected nature of zero-day threats that causes concern, especially because they may be used in targeted attacks and in the propagation of malicious code. As Symantec predicted in the Volume IX of the *Internet Security Threat Report*, a black market for zero-day vulnerabilities has emerged that has the potential to put them into the hands of criminals and other interested parties.[107]

In the second half of 2006, Symantec documented six zero-day vulnerabilities (figure 23). In the previous six-month period, Symantec documented 12 zero-day vulnerabilities. In the first half of 2006, only one zero-day vulnerability was documented.

[107] Symantec *Internet Security Threat Report*, Volume IX (March 2006): http://eval.symantec.com/mktginfo/enterprise/white_papers/ent-whitepaper_symantec_internet_security_threat_report_ix.pdf : p. 20

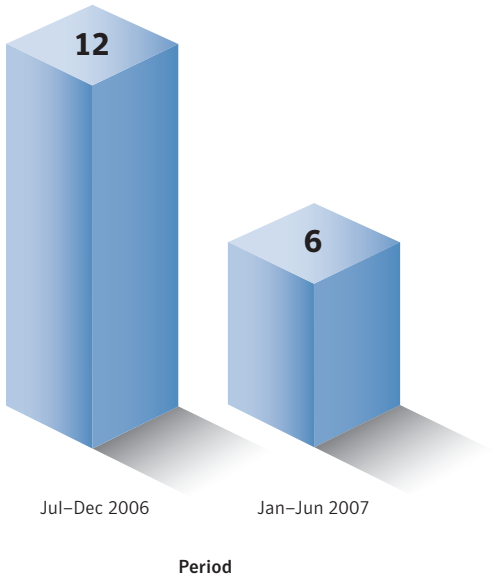

Figure 23. Zero-day vulnerabilities
Source: Symantec Corporation

Of the zero-day vulnerabilities documented during this period, three of the vulnerabilities affected Microsoft Office applications. This is a drop from the six zero-day vulnerabilities that affected Office in the second half of 2006. The number of zero-day Office vulnerabilities may have dropped due to measures taken by Microsoft to patch as many pending Office vulnerabilities as possible. During this reporting period, Microsoft also released an advisory describing the Microsoft Office Isolated Conversion Environment (MOICE) and File Block features and their applicability in mitigating zero-day vulnerabilities.[108] As the complexity of Microsoft Office contributes to the potential for vulnerabilities, these measures have been introduced by Microsoft to help users protect their computers against zero-day vulnerabilities.

In order to protect against zero-day vulnerabilities, Symantec recommends that administrators deploy network and host-based IDS/IPS systems as well as regularly updated antivirus software. Security vendors may provide rapid response to recently discovered zero-day vulnerabilities in the wild by developing and implementing new or updated IDS/IPS and antivirus signatures before a patch has been released by the affected vendor. Behavior-blocking solutions and heuristic signatures may also provide protection against zero-day vulnerabilities.

In addition, some IPS systems may provide further protection against memory corruption vulnerabilities in the form of address space layout randomization (ASLR),[109] and by making memory segments non-executable. These measures may complicate the exploitation of such vulnerabilities and make it more difficult for attack payloads to execute; however, this security measure may not protect all applications by default.

[108] http://www.microsoft.com/technet/security/advisory/937696.mspx
[109] Address space layout randomization is a security measure to complicate exploitation of some classes of vulnerabilities by randomizing the layout of process address space to make it less predictable to attackers.

Unpatched enterprise vendor vulnerabilities

In the previous volume of the Symantec *Internet Security Threat Report*, Symantec studied the vendor responsiveness to vulnerabilities and found that the majority of vulnerabilities were not being acknowledged, and therefore patched, by vendors.[110] That analysis provided insight into unpatched vulnerabilities without considering the size of the vendors affected.

This report expands on this study by examining the number of unpatched vulnerabilities affecting enterprise vendors whose applications are widely deployed and considered to be mission-critical in nature. The following enterprise vendors are reviewed in this section:

- Computer Associates
- Cisco
- EMC
- HP
- IBM
- McAfee
- Microsoft
- Oracle
- Sun
- Symantec

Unpatched vulnerabilities are publicly documented security issues that are not known to be patched by the vendor responsible for maintaining the affected application. Readers should note that the vulnerabilities discussed in this section were known to be unpatched at the time that the data was gathered for this report. They may have been patched in the meantime. There is also a likelihood that some of the vulnerabilities were patched by the vendor without a public announcement; in such cases there is insufficient publicly available information to label these issues as patched. It is also important to note that some unpatched vulnerabilities remain in this state because they affect unsupported products, or because the vendor has provided specific workarounds that address the vulnerability until a patch is available.

These vulnerabilities are a serious concern for enterprises because they cannot be resolved without applying best practices, workarounds, and mitigations. In many circumstances these measures will not provide complete protection against unpatched vulnerabilities.

In the first half of 2007, Symantec documented 90 unpatched enterprise vulnerabilities that were published during this period (table 5). Of these, 64 affected Microsoft, 13 affected Oracle, four affected Computer Associates, four affected HP, two affected IBM, two affected Symantec, and one vulnerability affected Sun. The rest of the vendors in the enterprise subset had no known vulnerabilities that were unpatched in this period.

[110] Symantec *Internet Security Threat Report*, Volume XI (March 2007): http://eval.symantec.com/mktginfo/enterprise/white_papers/ent-whitepaper_internet_security_threat_report_xi_03_2007.en-us.pdf : p. 42

Enterprise Vendors	Jan–Jun 2007	Jul–Dec 2006
Microsoft	64	75
Oracle	13	7
Computer Associates	4	0
HP	4	1
IBM	2	5
Symantec	2	1
Sun	1	3
McAfee	0	2

Table 5. Unpatched enterprise vulnerabilities by vendor
Source: Symantec Corporation

Of the enterprise vulnerabilities published in the second half of 2006, 94 were unpatched. 75 affected Microsoft, seven affected Oracle, five affected IBM, three affected Sun, two affected McAfee, one affected HP, and one vulnerability affected Symantec. No other enterprise vendors had vulnerabilities published during this period that remain unpatched.

Microsoft had the most unpatched vulnerabilities that were disclosed during the second half of 2006 and the first half of 2007. Many of the vulnerabilities in the sample set are considered lower severity, such as denial of service issues affecting client or desktop software. These issues may be considered a low priority by Microsoft. As a result, they may not typically be addressed in monthly security bulletins, but in service packs and other major version updates instead.

While it is likely that many of these vulnerabilities will have minimal impact on enterprises, some denial of service vulnerabilities have the potential for more severe effects such as code execution. Some vulnerabilities are prematurely thought to be limited to denial of service capabilities because the researcher has not completely investigated the vulnerability or because his or her skills are inadequate to conclusively determine the nature of the vulnerability.

The first half of 2007 did not show an improvement in the number of unpatched Oracle vulnerabilities over the second half of 2006. In addition to that, many vulnerabilities still remain unpatched from that period. In many cases, this may be due to lack of acknowledgement or correlation with publicly available vulnerability reports. When Oracle announces vulnerabilities, many of the issues are identified by an internal tracking number, but are not adequately mapped to other external vulnerability identifies such as the CVE dictionary.[111] This could cause many publicly known vulnerabilities to remain classified as unpatched because the vendor has not explicitly identified the vulnerabilities by their common names in security bulletins and product updates.

Recently, Oracle made improvements to their security reporting procedures, including providing pre-release notification for the security updates and including Common Vulnerability Scoring System ratings in their advisories.[112] The expectation is that these changes will have a positive effect on security reporting and vulnerability remediation. As a result, it is likely that fewer vulnerabilities will remain unpatched for extended periods of time.

[111] http://cve.mitre.org
[112] http://www.vnunet.com/articles/print/2172404

Browser plug-in vulnerabilities

For the first time, in this volume of the *Internet Security Threat Report*, Symantec is assessing vulnerabilities in browser plug-ins. These are technologies that run inside the Web browser and extend the browser's features. They can include plug-ins that permit additional multimedia content from Web pages to be rendered in the browser. It also includes execution environments that allow applications to be run inside the browser.

Many browsers include various plug-ins in their default installation and provide a framework to ease the installation of additional plug-ins. Plug-ins provide much of the expected or desired functionality of Web browsers. Some plug-ins may even be required to use public Web sites and/or an organization's internal sites. Browser plug-in vulnerabilities are implicated in some client-side attacks and present similar challenges to the enterprise.

This section examines vulnerabilities in the following browser plug-in technologies:

- Adobe Acrobat
- Adobe Flash®
- Apple QuickTime
- Microsoft ActiveX
- Microsoft Windows Media Player
- Mozilla browser extensions
- Opera widgets
- Sun Java

In the first half of 2007, Symantec documented 237 vulnerabilities affecting browser plug-ins (figure 24). Of these, 210 affected ActiveX components, 18 affected the Apple QuickTime plug-in, four affected the Sun Java plug-in, three affected extensions for Mozilla browsers, and two affected the Adobe Acrobat plug-in. Adobe Flash, Microsoft Windows Media Player, and Opera widgets were not affected by any browser plug-in vulnerabilities during this period.

There were 74 browser plug-in vulnerabilities documented during the second half of 2006. Of those, 43 vulnerabilities affected ActiveX components, eight affected Adobe Flash, eight affected the Apple QuickTime plug-in, seven affected the Adobe Acrobat plug-in, four affected the Sun Java plug-in, three affected Windows Media Player, and one was documented in Mozilla extensions. Opera widgets were not affected by any documented vulnerabilities in the second half of 2006.

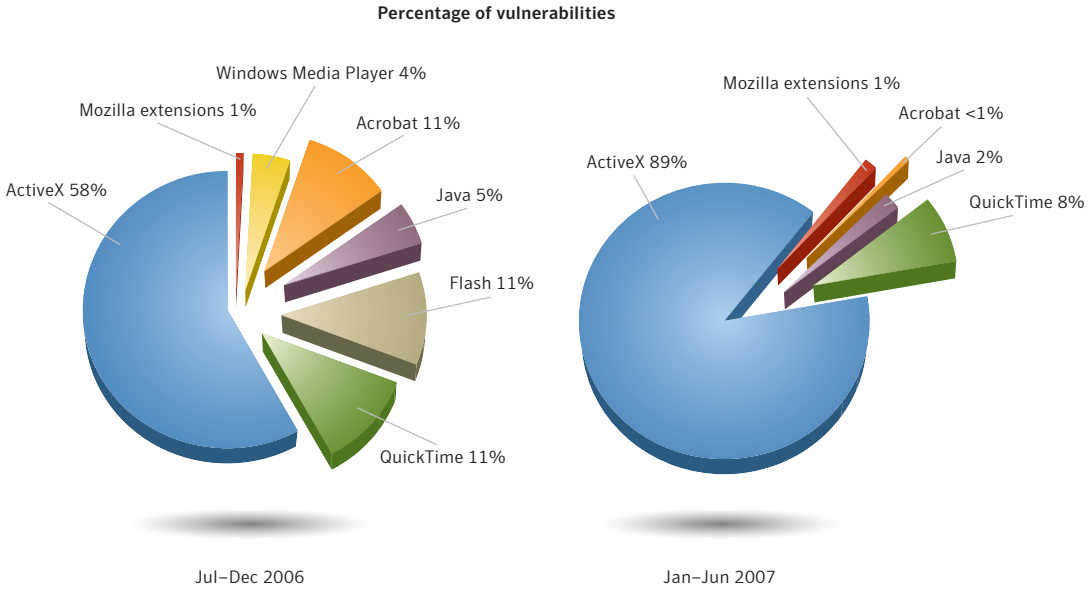

Figure 24. Browser plug-in vulnerabilities
Source: Symantec Corporation

The rise in browser plug-in vulnerabilities is indicative of an increasing focus on client-side vulnerabilities by both security researchers and attackers. The growth corresponds to an increase in the number of vulnerabilities in ActiveX components. This report expands on a previous Symantec study that observed the initial rise in ActiveX vulnerabilities.[113] That study determined that the use of fuzzers designed specifically to target insecure ActiveX components,[114] such as AxMan[115] and COMRaider,[116] has automated the discovery of these vulnerabilities. In addition, it is relatively easy to develop exploits for these types of vulnerabilities due to numerous examples of previous similar exploits that serve as a template.

These vulnerabilities affect a diverse group of vendors, including Microsoft, enterprise vendors, and smaller vendors. The sheer number of vulnerabilities gives attackers a wide range of potential targets. It should be noted that in addition to Windows Media Player, many of the affected ActiveX components may be included in default installations of Windows. Further, the installation and execution of ActiveX components is typically not evident to the user, while the removal of such components is difficult for the average end user. As a result, users may not be aware that they are prone to exploitation through vulnerable ActiveX components that have been installed on their computer.

Plug-in vulnerabilities have been the subject of exploit activity in the wild. For example, they were leveraged by many of the exploits employed by the MPack attack framework. In particular, MPack exploits a QuickTime vulnerability,[117] an issue in the WinZip ActiveX component,[118] and various other plug-in vulnerabilities such as the Microsoft WebView FolderIcon issue.[119] While some plug-ins may be specific to Internet Explorer, MPack also targets vulnerabilities in cross-browser plug-ins. This exposes users of alternate browsers on Windows by targeting shared weaknesses that are not necessarily dependant on how secure the browser itself is. The reliability and robustness of MPack implies that it benefited from professional development.

[113] http://www.symantec.com/enterprise/security_response/weblog/2007/01/a_sudden_rise_in_activex_vulne.html
[114] ActiveX components are a type of COM (Component Object Model) object that may provide a programming interface that is accessible through Internet Explorer. If exposed through Internet Explorer, attackers may exploit latent vulnerabilities in ActiveX components through malicious HTML content.
[115] http://www.metasploit.com/users/hdm/tools/axman
[116] http://labs.idefense.com/software/fuzzing.php#more_comraider
[117] http://www.securityfocus.com/bid/21829
[118] http://www.securityfocus.com/bid/21060
[119] http:/www.securityfocus.com/bid/19030

Client-side attacks have typically originated from questionable sources such as malicious Web sites or spam. As a result, best practices have advised end users to avoid this type of content. However, it appears that attackers are increasingly using legitimate and trusted sites as a basis for attacks. Symantec has observed that MPack includes functionality to deliver malicious payloads through legitimate Web sites that have been compromised.[120] In this scenario, it is necessary to exploit other unrelated vulnerabilities to deploy the attack framework to launch attacks against Web users. It integrates Web application vulnerabilities into attacks on the browser, whether directly or through plug-in and client-side vulnerabilities.

End users and administrators can use a number of measures to protect against the effects of vulnerabilities. IPS technologies can prevent exploitation of some browser plug-in vulnerabilities through signature- or behavior-based approaches in addition to ASLR. Antivirus software may also aid in protecting organizations from browser plug-in exploits through heuristic signatures.

While attacks are likely to originate from Web sites that are trusted as well as those that are not, Web browser security features can help reduce exposure to browser plug-in exploits, as can white-listing. Specifically, administrators and end users should actively maintain a white-list of trusted Web sites, and should disable individual plug-ins and scripting capabilities for all other sites. This will not prevent exploitation attempts from white-listed sites but may aid in preventing exploits from all other sites. Organizations can also implement a white-list policy at the network perimeter to regulate outgoing access by end users. Content filtering may also be employed to strip malicious content from trusted and untrusted sites.

Vulnerabilities—protection and mitigation

In addition to the specific steps required to protect against the vulnerabilities discussed in this section, there are general steps that should be taken to protect against the exploitation of vulnerabilities. Administrators should employ a good asset management system to track what assets are deployed on the network and to determine which ones may be affected by the discovery of new vulnerabilities. Vulnerability management technologies should also be used to detect known vulnerabilities in deployed assets. Administrators should monitor vulnerability mailing lists and security Web sites to keep abreast of new vulnerabilities in Web applications.

Symantec recommends that administrators employ vulnerability assessment services, a vulnerability management solution, and vulnerability assessment tools to evaluate the security posture of the enterprise. Unpatched vulnerabilities should be identified by administrators, and assessed and mitigated according to the risk they present. Where possible, problematic applications with many unpatched vulnerabilities should be removed or isolated. IPS systems can aid in detecting known attacks against such applications.

[120] http://www.symantec.com/enterprise/security_response/weblog/2007/06/mpack_the_strange_case_of_the.html

Enterprises should subscribe to a vulnerability alerting service in order to be notified of new vulnerabilities. They should also manage their Web-based assets carefully. If they are developing Web applications in-house, developers should be educated about secure development practices, such as the Security Development Lifecycle and threat modeling.[121] If possible, all Web applications should be audited for security prior to deployment. Web application security solutions and a number of products and services are available to detect and prevent attacks against these applications.

When deploying applications, administrators should ensure that secure, up-to-date versions are used, and that applications are properly configured to avoid the exploitation of latent vulnerabilities. Symantec recommends the use of secure shared components that have been audited for common Web application vulnerabilities. As much as possible, enterprises are advised to avoid deploying products that are not regularly maintained or that are not supported by the vendor.

[121] The Security Development Lifecycle is a development paradigm that incorporates security at every stage from the initial architecture to programming, and in the quality assurance/testing phases. Threat modeling is a security auditing methodology that involves formally identifying and mapping out all possible attack vectors for an application.

Malicious Code Trends

Symantec gathers malicious code data from over 120 million desktops that have deployed Symantec's antivirus products in consumer and corporate environments. The Symantec Digital Immune System and Scan and Deliver technologies allow customers to automate this reporting process. This discussion is based on malicious code samples reported to Symantec for analysis between January 1 and June 30, 2007.

In previous editions of the Symantec *Internet Security Threat Report*, the number and volume of threats analyzed were based upon the number of malicious code reports received from enterprise and home users. This report will also examine malicious code according to potential infections. This allows Symantec to determine which malicious code sample was attempting to infect computers and the number of potential infections worldwide.

This discussion will include any prevention and mitigation measures that might be relevant to the particular threats being discussed. However, Symantec recommends that certain best security practices always be followed to protect against malicious code infection. Administrators should keep patch levels up to date, especially on computers that host public services and applications—such as HTTP, FTP, SMTP, and DNS servers—and that are accessible through a firewall or placed in a DMZ. For organizations and businesses, email servers should be configured to only allow file attachment types that are required for business needs and to not accept email that appears to come from within the company but originates from external sources. Additionally, Symantec recommends that ingress and egress filtering be put in place on perimeter devices to prevent unwanted activity.

End users should employ defense-in-depth strategies, including the deployment of antivirus software and a personal firewall. Users should update antivirus definitions regularly. They should also ensure that all desktop, laptop, and server computers are updated with all necessary security patches from their software vendors. They should never view, open, or execute any email attachment unless it is expected and comes from a trusted source, and unless the purpose of the attachment is known.

Malicious Code Trends Highlights

The following section will offer highlights of the malicious code trends that Symantec observed during this period. Following this overview, the *Internet Security Threat Report* will discuss selected metrics in greater depth, providing analysis and discussion of the trends indicated by the data.

- Of the top ten new malicious code families detected in the first six months of 2007, four were Trojans, three were viruses, one was a worm, and two were worms with a virus component.
- In the first half of 2007, 212,101 new malicious code threats were reported to Symantec. This is a 185 percent increase over the second half of 2006.
- During the first half of 2007, Trojans made up 54 percent of the volume of the top 50 malicious code reports, an increase over the 45 percent reported in the final six months of 2006.
- When measured by potential infections, Trojans accounted for 73 percent of the top 50 malicious code samples, up from 60 percent in the previous period.

- 43 percent of worms reported this period originated in the Europe, Middle East, and Africa (EMEA) region.
- North America accounted for 44 percent of Trojans reported this period.
- Threats to confidential information made up 65 percent of the top 50 potential malicious code infections reported to Symantec.
- Of all confidential information threats detected this period, 88 percent had a keystroke logging component and 88 percent had remote access capabilities, an increase from 76 percent and 87 percent, respectively, over the previous period.
- Forty-six percent of malicious code that propagated did so over SMTP, making it the most commonly used propagation mechanism.
- During the first half of 2007, 18 percent of the 1,509 documented malicious code instances exploited vulnerabilities.
- Thirty-five percent of infected computers reported more than one infection in the first half of 2007.
- Eight of the top ten staged downloaders this period were Trojans and two were worms.
- Seven of the top ten downloaded components were Trojans and three were back doors.
- Malicious code that targets online games made up five percent of the top 50 potential malicious code infections.
- Lineage and World of Warcraft were the two most frequently targeted online games in the first half of 2007.

Malicious Code Trends Discussion

This section will discuss selected "Malicious Code Trends" metrics in greater depth, providing analysis and discussion of the trends indicated by the data. The following metrics will be discussed:

- Top ten new malicious code families
- New malicious code threats
- Malicious code types
- Geolocation by type
- Threats to confidential information
- Propagation mechanisms
- Malicious code that exploits vulnerabilities
- Percentage of computers with multiple infections
- Staged downloaders—multiple infections by type
- Malicious code targeting online gaming
- Malicious code—prevention and mitigation

Top ten new malicious code families

Of the top ten new malicious code families detected in the first six months of 2007, four were Trojans, three were viruses, one was a worm, and two were worms with a virus component (table 6). One of the Trojans also had back door capabilities. This indicates that attackers may be moving towards using Trojans as a means of installing malicious code on computers. This is indicative of multistaged attacks, in which an initial compromise takes place that is not intended to perform malicious activities immediately, but that is used to facilitate the launch of subsequent attack activity. Symantec believes that multistaged attacks are becoming more common, as attackers adopt new tactics to circumvent effective security measures that have evolved to prevent previous attack methods.

As Trojans do not propagate, they allow attackers to perform targeted attacks without drawing attention to themselves. Worms, on the other hand, propagate by sending themselves in high volumes of email messages or by attacking other computers, thereby increasing the likelihood of being noticed by network administrators who can take immediate action. A Trojan that is installed when a user visits a malicious Web site or downloads and opens a malicious file is much more likely to escape notice, as there will be no high-volume traffic associated with it. This degree of stealth increases the Trojan's effectiveness. The longer a threat remains undiscovered in the wild, the more opportunity it has to compromise computers before measures can be taken to protect against it. Furthermore, its ability to steal information increases the longer it remains undetected on a compromised computer.

Rank	Sample	Type	Vectors	Impacts/Features
1	Peacomm	Trojan	Spam/Mixor.Q	Creates an encrypted peer-to-peer network and downloads other threats
2	Whybo	Virus	File Sharing	Downloads and executes other files
3	Metajuan	Trojan	N/A	Downloads other threats and displays ads
4	Anivip	Virus	File Sharing/ Remote Vulnerability	Downloads other threats
5	Kakavex	Virus	File Sharing	Steals credit card information
6	Pandex	Trojan	N/A	Gathers email addresses and relays spam
7	Fakerecy	Worm	File Sharing	Copies itself to all fixed, removable, and network drives
8	Validin	Worm/Virus	File Sharing	Downloads other threats
9	Fubalca	Worm/Virus	File Sharing	Downloads other threats
10	Mespam	Trojan	Peacomm	Sends instant messages containing a malicious URL

Table 6. Top ten new malicious code families
Source: Symantec Corporation

The most widely reported new malicious code family during this reporting period was the Peacomm Trojan,[122] also known as the Storm Trojan. This Trojan was spammed in high volumes by the Mixor.Q worm,[123] which prompted Symantec to classify it as a Category 3 threat in January 2007.[124]

[122] http://www.symantec.com/security_response/writeup.jsp?docid=2007-011917-1403-99
[123] http://www.symantec.com/security_response/writeup.jsp?docid=2006-122917-0740-99
[124] A Category 3 threat is a malicious code sample that is considered a moderate threat. It is either currently spreading among computer users but reasonably harmless and easy to contain, or has not been released into the wild but is potentially dangerous and difficult to contain.

When Peacomm installs itself on a computer, it attempts to hide itself using rootkit techniques.[125] It also contains a list of other compromised computers that it uses to build an encrypted network of peers. This is similar to a bot network; however, rather than using IRC to communicate, as bot networks traditionally do, it uses the Overnet peer-to-peer protocol in order to make the network more resilient since this approach has no single point of failure.[126]

Peacomm listens for commands passed through its peer-to-peer network and then downloads and installs other files, such as the Mespam[127] and Abwiz.F Trojans.[128] This can be of particular concern, since a Trojan like Abwiz.F can send confidential information to the remote attacker and relay spam.

The Whybo virus was the second most common new malicious code family in the first half of 2007.[129] This virus infects portable executable files on all drives from C to Z on the compromised computer. It also retrieves an encrypted file from a remote computer and executes it. Plus, it closes open windows with certain strings in their titles, some of which are related to security applications. Interestingly, the virus matches these strings in both English and Chinese, indicating that it was likely written by someone familiar with both languages. It may also indicate that Whybo was intended to particularly target users in China.

The Metajuan Trojan was the third most frequently reported new malicious code family this period.[130] This Trojan may be installed by other malicious code samples or installed by Web pages that are designed to exploit Internet Explorer vulnerabilities. This means that the user will be compromised by visiting a malicious Web site rather than receiving the Trojan through email. This represents a trend in which attackers are relying upon users to retrieve threats instead of sending the threat directly to potential victims.

Metajuan also illustrates the current trend towards multistaged attacks. Once installed, the Trojan contacts a remote Web site and can download and execute other malicious files on the compromised computer. Metajuan may also display advertisements when the user visits certain Web pages.

Kakavex was the fifth most common new malicious code family in the first half of 2007.[131] This virus is notable because it may represent the beginning of an interesting trend. Traditionally, most viruses simply infect executable files and perform some form of damaging action. However, in addition to infecting files, the Kakavex virus also attempts to steal credit card information. The virus monitors Internet usage on the infected computer and, under certain circumstances, may display a dialogue box prompting the user for his or her credit card information. The information is then sent to a remote Web site.

This virus shows that identity thieves appear to be expanding into new territory to steal personal information. In the past they mainly used back doors and Trojans to steal this kind of information; however, Kakavex indicates that they are now using viruses to do the same thing, thereby expanding the number of tools available to them for this objective.

[125] Rootkit techniques are used by malicious code to hide their presence on a compromised computer.
[126] Overnet is a decentralized peer-to-peer file-sharing protocol. It was taken down due to legal action in September 2006, but due to its decentralized nature, clients are still able to function.
[127] http://www.symantec.com/security_response/writeup.jsp?docid=2007-020915-2914-99
[128] http://www.symantec.com/security_response/writeup.jsp?docid=2006-032311-1146-99
[129] http://www.symantec.com/security_response/writeup.jsp?docid=2007-040316-2416-99
[130] http://www.symantec.com/security_response/writeup.jsp?docid=2007-030112-0714-99
[131] http://www.symantec.com/security_response/writeup.jsp?docid=2007-011014-1759-99

New malicious code threats

The number of new malicious code threats detected by Symantec in a given reporting period allows administrators and users to keep track of the productivity of malicious code writers in a given period. These malicious code samples are collected through submissions received from Symantec customers as well as from Symantec honeypot computers.[132] Periods in which large amounts of new malicious code are created require frequent updating of antivirus signatures, as well as the implementation of other security measures, such as patching against Web browser vulnerabilities that are frequently exploited to install malicious code on computers.

In the first six months of 2007, Symantec detected 212,101 new malicious code threats (figure 25). This is a 185 percent increase over the previous period when 74,482 new threats were detected and a 318 percent increase over the first half of 2006. This brings the total amount of threats identified by Symantec to 622,500 as of the end of June 2007. This means that more than one third of all malicious code threats currently detected were created in the first six months of 2007.

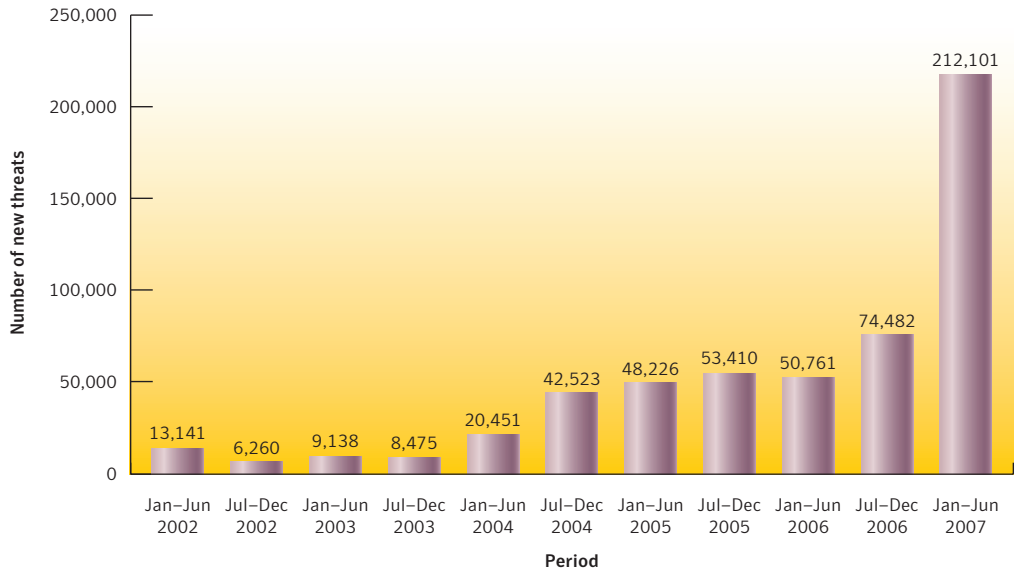

Figure 25. New malicious code threats
Source: Symantec Corporation

[132] A honeypot is an Internet-connected system that acts as a decoy, allowing an attacker to enter the system so that the attacker's behavior inside the compromised system can be observed.

The increase in threats this period can mainly be attributed to an increase in new Trojans, including staged downloaders, which consist of a small Trojan that downloads and installs other malicious code on a computer. The initial Trojan is frequently written for a specific purpose or target. For example, the initial stage may be installed by a Web page that exploits a browser vulnerability. In some cases, the downloader may be written to only download and install a particular file from a specific location. To avoid being noticed, this Trojan is usually quite small in size to avoid detection and establish a "beachhead" for subsequent infections. The main functionality of a staged downloader system is contained in the second or possibly third stage.

The high quantity of production of these downloaders demonstrates the need to ensure that antivirus signatures are kept up-to-date on a regular basis. Since signatures are created in response to new threats in the wild, it is vital that end users and enterprises maintain the most current antivirus definitions in order to protect against rapidly launched new threats.

Malicious code types

During the first half of 2007, Trojans made up 54 percent of the volume of the top 50 malicious code reports, an increase over the 45 percent reported in the final six months of 2006 (figure 26). While part of this increase can be attributed to the success of the Peacomm Trojan, there were also a wide variety of other Trojans present in the top 50 malicious code reports. As previously mentioned, Trojans are likely gaining prominence because they generate a low volume of traffic compared to network and mass-mailing worms. As a result, they are less likely to draw the attention of higher-profile threats. Furthermore, malicious code writers may be turning to Trojans because network perimeter defenses and desktop firewalls, neither of which affects Trojans, make it harder for network worms to propagate widely.

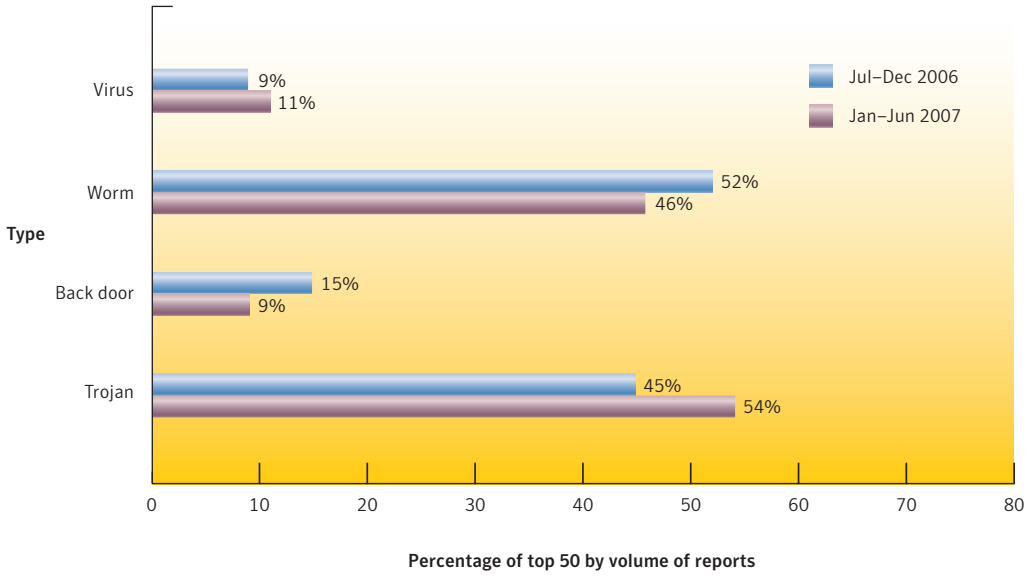

Figure 26. Malicious code types by volume
Source: Symantec Corporation

Trojans may also be gaining popularity because they are better suited to meet the objectives of attackers. As was stated previously, Trojans are likely able to reside on infected computers for longer periods of time. This allows them to remain active on the computer longer, enhancing the opportunity to gather confidential information, download malicious components for subsequent attacks, and/or cause more damage. For example, the Vundo Trojan installs adware on a compromised computer.[133] Variants of the Adclicker Trojan can be used to generate traffic to Web sites in order to increase revenue from banner ads,[134] a practice commonly referred to as click fraud.

Additionally, other Trojans can be used to relay spam email or in phishing attacks. For instance, the Flush Trojan modifies the DNS settings on a compromised computer,[135] which can cause the user's Web browser to be redirected to a phishing site when he or she attempts to connect to an online banking site. The high volume of these Trojans in the top 50 malicious code reports demonstrates the popularity among attackers of utilizing malicious code to generate revenue.

During the first six months of 2007, worms made up 46 percent of the volume of the top 50 malicious code samples reported to Symantec, down from 52 percent in the previous period. This is a continuation of a downward trend in worm reports over the last year, which has been caused by a combination of a decrease in the volume of worms, as well as an increase in the volume of Trojans and viruses.

Worm numbers in the first half of 2007 were bolstered by Blackmal.E[136] and several variants of Mytob.[137] While reports of these worms continue to persist, they are nowhere near the levels of a year ago, when they made up 75 percent of the volume of the top 50 malicious code samples. The Mixor.Q worm, which was discovered at the end of 2006, was reported in significant numbers this period. However, this worm emailed copies of the Peacomm Trojan in significant numbers, which increased the volume of Trojans reported this period.

Part of the reason for the decline in worms has been due to security measures put in place, such as email attachment blocking at the SMTP gateway and blocking ports used by peer-to-peer file-sharing applications to prevent their propagation within the enterprise's network space. As a result, malicious code authors are likely looking for additional mechanisms to allow their creations to propagate, such as including a viral propagation component in traditionally non-viral malicious code.

Viruses made up 11 percent of the volume of top 50 malicious code reports in the first six months of 2007, a slight increase over the nine percent in the previous six-month period. The increase in viruses is related to a rise in the number of worms that also employ a file infection component, which causes them to also be classified as viruses. One example of this is Looked.BK,[138] which infects executable files in local drives and network shares.

[133] http://www.symantec.com/security_response/writeup.jsp?docid=2004-112111-3912-99
[134] http://www.symantec.com/security_response/writeup.jsp?docid=2002-091214-5754-99
[135] http://www.symantec.com/security_response/writeup.jsp?docid=2005-030413-5303-99
[136] http://www.symantec.com/security_response/writeup.jsp?docid=2006-011712-2537-99
[137] http://www.symantec.com/security_response/writeup.jsp?docid=2005-022614-4627-99
[138] http://www.symantec.com/security_response/writeup.jsp?docid=2006-112813-0222-99

In addition to assessing malicious code according to the volume of unique samples reported to Symantec, the *Internet Security Threat Report* assesses it according to the number of potential infections. This is an important distinction. In some cases, a threat that may be widely reported may not cause a large number of potential infections and vice versa.

The distinction between malicious code reports and infections is well illustrated by comparing worm and Trojan activity. While worms made up 46 percent the volume of the top 50 malicious code reports in the first half of 2007, they caused only 22 percent of potential infections (figure 27). The main reason for this is that mass-mailing worms generate a significant number of email messages to which they attach their malicious code. Each message that is detected will generate a malicious code report. Because of the high volume of email that one worm can generate, a single infection can result in many reports. However, once a malicious code sample is detected, antivirus signatures are quickly developed that can protect against subsequent potential infections by that sample. So, only a small percentage of the high volume of email messages will result in potential infections.

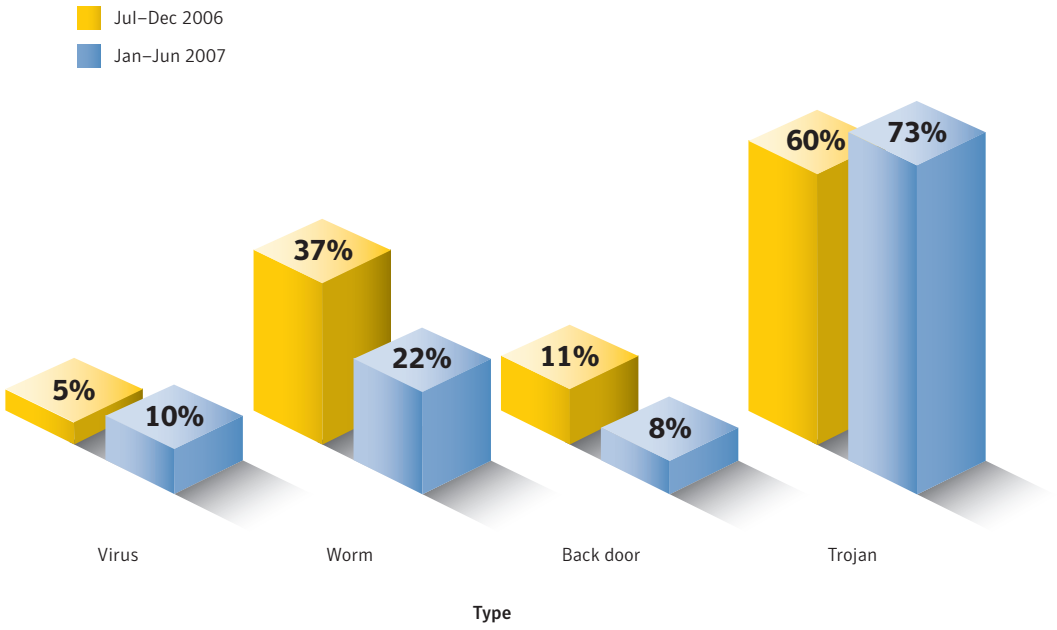

Figure 27. Malicious code types by potential infections
Source: Symantec Corporation

As was mentioned previously in this section, Trojans made up 54 percent of the volume of the top 50 malicious code reports the first half of 2007. In terms of potential infections, Trojan activity represented 73 percent of malicious code activity during this period, up from 60 percent in the second half of 2006. At the same time, potential infections caused by worms declined from 37 percent in the second half of 2006 to 22 percent in the first six months of 2007. It stands to reason that since users are seeing fewer worms—for example, fewer email messages from mass-mailing worms in their inboxes—as indicated by the decline in reports above, they are also less likely to be infected by a worm. The number of unique Trojans and worms in the top 50 potential infections is close, at 22 and 20 respectively, but the volume of Trojans far outweighs the volume of worms in the period.

Viruses experienced significant growth in potential infections during the first six months of 2007. While viruses increased slightly in the volume of the top 50 malicious code reports, the number of potential infections doubled from five percent in the previous period to ten percent in the current period. As previously stated, this is likely a result of new worms that also employ a viral component in order to propagate.

Geolocation by type

For the first time, in this edition of the *Internet Security Threat Report*, Symantec is examining the top regions reporting potential malicious code infections, as well as the types of malicious code causing potential infections in each region. The increasing regionalization of threats can cause differences between the types of malicious code being observed from one area to the next. For example, threats may use certain languages or localized events as part of their social engineering techniques. Threats that steal confidential information can also be tailored to steal information that is more common in some countries than in others. Trojans that steal account information for Brazilian banks are quite common in the Latin America region, while malicious code that steals online gaming account information is most frequently observed in the Asia-Pacific and Japan region. Because of the different propagation mechanisms used by different malicious code types, and the different effects that each malicious code type may have, the geographic distribution of malicious code can illustrate how network administrators in different regions can best increase the focus of their security efforts.

Between January and June of 2007, 44 percent of Trojans were reported from North America, while 37 percent were reported from the EMEA region (figure 28). This is significantly higher than the 15 percent reported from the Asia-Pacific and Japan (APJ) region and the four percent from Latin America.

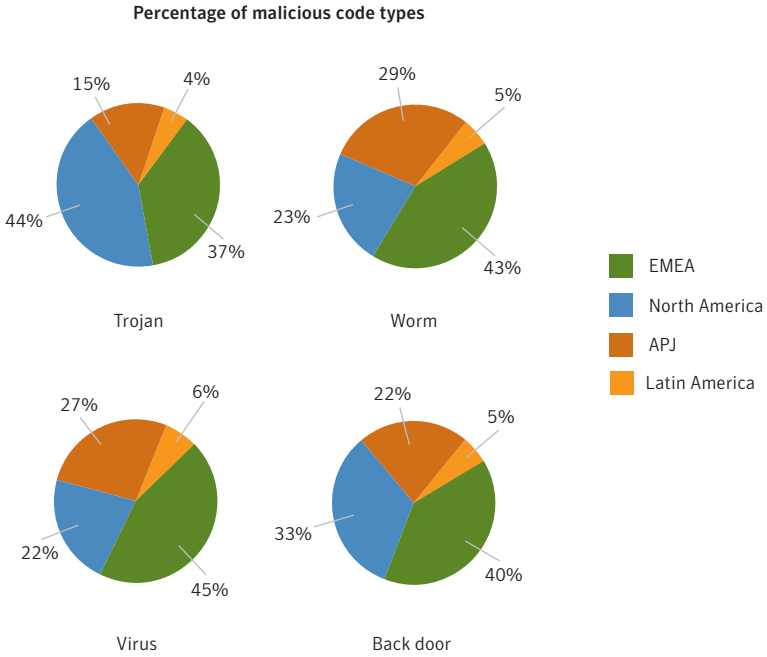

Figure 28. Location of malicious code by type
Source: Symantec Corporation

The concentration of Trojans in North America may be indicative of enterprises and ISPs taking more active steps to prevent the propagation of worms. Steps include more aggressive blocking and filtering of email attachments at the email gateway to prevent the propagation of mass-mailing worms, and port blocking to prevent the spread of network worms. The prevalence of Trojans in North America could be reflective of the resultant drop in network worms in the region. On the other hand, it could reflect a conscious decision by attackers to move towards Trojans in reaction to the success of tactics that have successfully thwarted worm attacks.

As discussed in the "Malicious code that exploits vulnerabilities" section below, many Trojans are now being installed by Web pages that exploit vulnerabilities. This indicates that users and enterprises in regions with higher Trojan concentrations should ensure that their Web browsers, as well as related components and plug-ins, are patched for any potential vulnerabilities.

During this period, EMEA accounted for 43 percent of global potential infections caused by worms. This was followed by the APJ region, which accounted for 29 percent of potential worm infections. North America only accounted for 23 percent of reported worms this period. This may indicate that North American ISPs are implementing more rigid port blocking to limit the spread of network worms, as well as antivirus filtering at the email gateway to limit mass-mailing worms.

Some worms use region-specific subject lines and text in their email messages. For example, the Rontokbro worm's email messages are in Indonesian.[139] However, this worm was seen more in India than in any other country. There is a great deal of commerce between India and Indonesia,[140] which means that it is highly likely that many enterprise users in Indonesia communicate with counterparts in India by email. Since Rontokbro sends its email messages to all the addresses it gathers from files on a compromised computer, it stands to reason that this worm was sent to many Indian users from business contacts in Indonesia. Rontokbro was also one of the top ten malicious code samples resulting in potential infections in the EMEA region.

The EMEA region accounted for the highest percentage of viruses this period, with 45 percent of the total. The APJ and North America regions accounted for 27 and 22 percent of viruses respectively, while Latin America only accounted for six percent.

As is noted in the "Malicious code types" section of this report, many worms are incorporating a viral component that causes them to be classified as both worms and viruses. Many of the worms causing potential infections in EMEA also employ a viral component, which explains why this region also accounts for the greatest percentage of viruses and worms this period.

Potential infections caused by back doors were most frequently reported from the EMEA region, which accounted for 40 percent of all back doors worldwide. North America accounted for 33 percent of potential back door infections in the first half of 2007, while APJ accounted for 22 percent and Latin America accounted for five percent. It is important to note that while the regional percentages of potential back door infections show a fairly wide variance during this period, the worldwide volume of back door threats this period was significantly lower than Trojans and worms. As a result, the percentage variance between regions actually represents a much smaller difference in raw numbers than the percentage differences between worms and Trojans.

[139] http://www.symantec.com/security_response/writeup.jsp?docid=2005-092311-2608-99
[140] http://www.hindu.com/2005/11/24/stories/2005112405871200.htm

Threats to confidential information

Some malicious code programs are designed specifically to expose confidential information that is stored on an infected computer. These threats may expose sensitive data such as system information, confidential files and documents, or logon credentials. Some malicious code threats, such as back doors, can give a remote attacker complete control over a compromised computer.

Threats to confidential information are a particular concern because of their potential for use in criminal activities. With the widespread use of online shopping and Internet banking, compromises of this nature can result in significant financial loss, particularly if credit card information or banking details are exposed.

Within the enterprise, exposure of confidential information can lead to significant data leakage. If it involves customer-related data—such as credit card information—this can severely undermine customer confidence as well as violate local laws. Sensitive corporate information, including financial details, business plans, and proprietary technologies, could also be leaked from compromised computers. It should be noted that threats that expose confidential information may employ more than one method to do so; as a result, cumulative percentages discussed in this metric may exceed 100 percent.

In the first six months of 2007, threats to confidential information made up 65 percent of potential infections by the top 50 malicious code samples (figure 29). This is an increase over the 53 percent of potential infections in the second half of 2006.

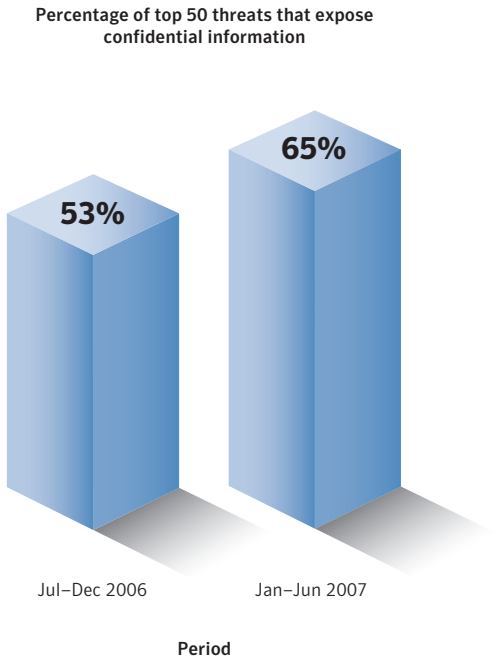

Figure 29. Threats to confidential information by volume
Source: Symantec Corporation

Malicious code can expose confidential information in a variety of ways. The most common method is by allowing remote access to the compromised computer through a back door. In this method, the attacker typically uses a specialized application to connect to the compromised computer. He or she can then perform numerous actions such as taking screenshots, changing configuration settings, and uploading, downloading, or deleting files.

In this reporting period, 88 percent of confidential information threats had a remote access component (figure 30). Remote access threats made up 87 percent of confidential information threats in the second half of 2006. Back doors typically require a two-way communication channel between the attacker and the compromised computer in order to access unauthorized information. As such, they may be less efficient than an automated mechanism, such as a keystroke logger. This may indicate why threats that allow remote access only increased marginally this period while other information exposure types increased more significantly.

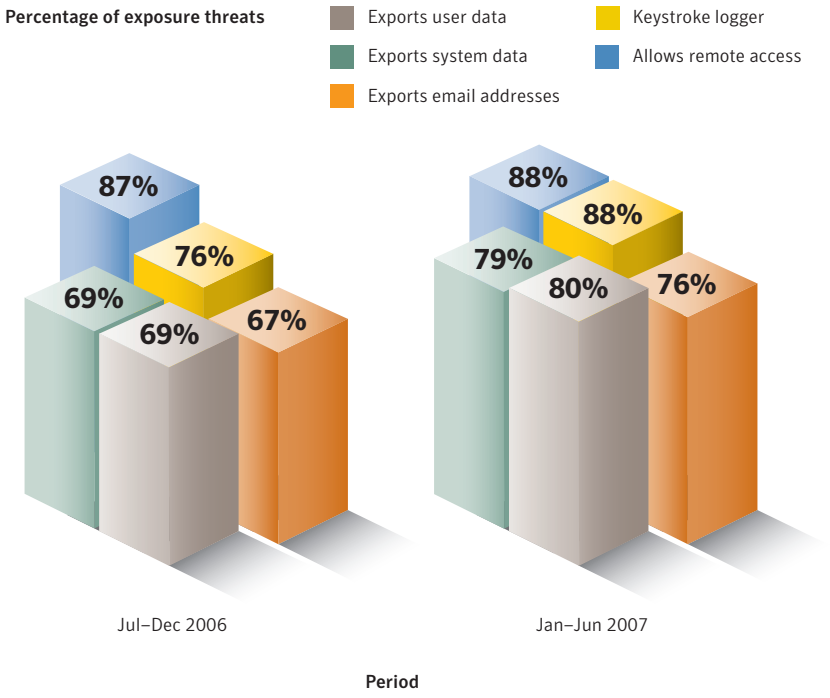

Figure 30. Threats to confidential information by type
Source: Symantec Corporation

Confidential information threats with keystroke logging capability made up 88 percent of threats to confidential information, up from 76 percent in the second half of last year. A keystroke logger records keystrokes on a compromised computer and either emails the log to the attacker or uploads it to a Web site under the attacker's control. This makes it easier for the attacker to gather confidential information from a large number of compromised computers than if he or she had to manually connect to back doors installed on various computers. The attacker can use these logs to find the user's credentials for different types of accounts, such as online banking and trading accounts, as well as ISP accounts. The attacker can then use this information as a stepping stone to launch further attacks.

Threats that could be employed to export user data accounted for 80 percent of confidential information threats during the first six months of 2007, up from 69 percent in the previous reporting period. Furthermore, in the first half of 2007, 79 percent of threats to confidential information could be used to export system data, compared to 69 percent in the second half of 2006. These forms of data leakage can be used to steal a user's identity or launch further attacks. Attackers with access to the user's personal and system data can use it to craft a more targeted social engineering attack tailored to that particular user.

Organizations can take several steps to limit the exposure of confidential information by successful intrusions. Encrypting sensitive data that is stored in databases will limit an attacker's ability to view and/or use the data. However, this step will require that sufficient computing resources be made available, as encrypting and decrypting the data for business use consumes processing cycles on servers. Furthermore, encrypting stored data will not protect against so-called man-in-the-middle attacks that intercept data before it is encrypted.[141] As a result, data should always be transmitted through secure channels such as SSH, SSL, and IPSec.

Propagation mechanisms

Worms and viruses use various means to transfer themselves, or propagate, from one computer to another. These means are collectively referred to as propagation mechanisms. This section will assess some of the propagation mechanisms used by malicious code samples that were reported to Symantec in the first half of 2007. It will assess these samples according to the percentage of potential infections. Readers should note that some malicious code samples use more than one mechanism to propagate. As a result, cumulative percentages presented in this discussion may exceed 100 percent.

Due to some methodological changes that Symantec made for this reporting period, this volume of the *Internet Security Threat Report* is able to examine propagation mechanisms with increased specificity. For example, where possible, the specific peer-to-peer protocols employed as propagation mechanisms have been identified. This will allow administrators to look at more specific port blocking and protocol filtering based upon the specific propagation mechanisms being discussed. It is also important to note that, due to this change, any comparisons to previous reporting periods would not be valid; therefore, they have not been presented here.

In the second half of 2007, 46 percent of malicious code that propagated did so in email attachments (table 7). This is not surprising, given the pervasive use of email. However, as noted in the "Top ten new malicious code families" section of this report, malicious code authors seem to be diversifying their propagation mechanisms by combining worms with a viral file-infection component.

To limit the propagation of email-borne threats, administrators should ensure that all email attachments are scanned at the gateway. Additionally, all executable files originating from external sources, such as email attachments or downloaded from Web sites should be treated as suspicious. All executable files should be checked by antivirus scanners using the most current definitions.

[141] A "man-in-the-middle attack" is a form of attack in which a third party intercepts communications between two computers. The "man in the middle" captures the data, but still relays it to the intended destination to avoid detection. This can allow the attacker to intercept communications on a secure or encrypted channel.

Rank	Propagation Mechanism	Percentage of Threats
1	File Transfer/Email Attachment	46%
2	File Transfer/CIFS	24%
3	File Sharing/Peer-to-Peer	22%
4	File Sharing/Executables	22%
5	File Sharing/Peer-to-Peer/Kazaa	18%
6	Remotely Exploitable Vulnerability	18%
7	File Sharing/Peer-to-Peer/Morpheus	15%
8	File Sharing/Peer-to-Peer/eDonkey	15%
9	File Sharing/Peer-to-Peer/Winny	5%
10	Backdoor/Kuang2	3%

Table 7. Propagation mechanisms
Source: Symantec Corporation

Of the malicious code that propagated during the first half of 2007, 24 percent did so by the Common Internet File Sharing (CIFS) protocol.[142] Malicious code samples such as Fujacks.E[143] and variants of the Looked[144] family both propagated in significant numbers this period by copying themselves to CIFS shares with weak password protection. Both of these worms also contain a viral component to infect portable executable files. Since they try to infect files on both local and mapped network drives, they effectively use this propagation mechanism multiple times.

This propagation mechanism can be threatening to organizations because file servers use CIFS to give users access to their file shares. If a computer with access to a file server becomes infected by a threat that propagates through CIFS, it could spread to the file server. Since multiple computers within a corporation likely access the same file server, this could facilitate the rapid propagation of the threat within the enterprise.

To protect against threats that use the CIFS protocol to propagate, all shares should be protected with strong passwords, and only users who require the resources should be given access to them. If other users do not need to write to a share, they should only be given “read” permissions. This will prevent malicious code from copying itself to the shared directory or modifying shared files. Finally, CIFS shares should not be exposed to the Internet. Blocking TCP port 445 at the network boundary will help to protect against threats that propagate using CIFS.

Malicious code using peer-to-peer (P2P) protocols to propagate accounted for 22 percent of all potential infections this period. These samples typically do not attempt to use a specific P2P protocol to propagate; rather they copy themselves to all folders on a computer containing the string “shar”. P2P applications commonly create folders containing the word “share”—such as “shared folder”—so these malicious code samples will successfully propagate through many of them.

Four specific P2P protocols were commonly used by malicious code to propagate during the first six months of 2007. The Kazaa file-sharing service was used by 18 percent of malicious code samples that propagated, while Morpheus and eDonkey were each used by 15 percent. Finally, the Winny protocol was used by five percent of propagating malicious code this period.

[142] CIFS is a file sharing protocol that allows files and other resources on a computer to be shared with other computers across the Internet. One or more directories on a computer can be shared to allow other computers to access the files within.
[143] http://www.symantec.com/security_response/writeup.jsp?docid=2007-010509-0134-99
[144] http://www.symantec.com/security_response/writeup.jsp?docid=2004-121709-0657-99

Since P2P applications are typically not permitted on corporate networks, any P2P clients are likely installed without the knowledge or consent of network administrators. Enterprises should take measures to prevent P2P clients from being installed on any computers on the network. They should also block any ports used by these applications at the network boundary. End users who download files from P2P networks should scan all such files with a regularly updated antivirus product.

Malicious code that exploits vulnerabilities

The exploitation of vulnerabilities as a means of malicious code propagation is an ongoing concern for enterprises. This section of the *Internet Security Threat Report* will assess the proportion of malicious code that exploits vulnerabilities. This can provide some insight into how popular are vulnerabilities among malicious code authors when developing malicious code and variants thereof. The number of malicious code samples exploiting vulnerabilities gives administrators an indication of the need to apply patches in a timely manner.

During the first half of 2007, 18 percent of the 1,509 documented malicious code instances exploited vulnerabilities (figure 31).[145] This is lower than the 23 percent of the 1,318 malicious code instances documented in the second half of 2006. While the number of new samples exploiting vulnerabilities declined in the current reporting period, this method of propagation remains effective, as is illustrated by its presence in the top ten propagation mechanisms (table 7).

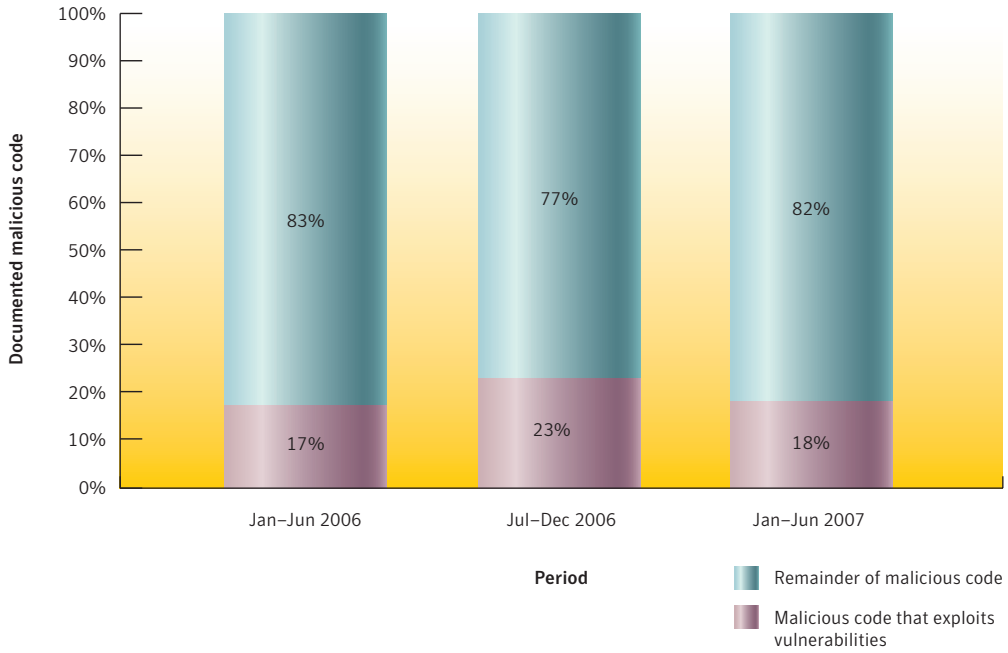

Figure 31. Malicious code that exploits vulnerabilities
Source: Symantec Corporation

[145] It should be noted that the number of documented malicious code instances differs from the number of malicious code submissions. Documented malicious code instances are those that have been analyzed and documented within the Symantec malicious code database.

The recent decline in the number of malicious code samples exploiting vulnerabilities is related to the drop in the number of zero-day vulnerabilities during this period. As discussed in the "Vulnerability Trends" section of this report, the number of new zero-day vulnerabilities documented decreased from 12 in the second half of 2006 to six in the first six months of 2007. Since a patch does not exist for a zero-day vulnerability, it is an effective way for malicious code to be installed on a vulnerable computer. As a result, the number of new zero-day vulnerabilities in a period can have a direct effect on the number of threats that are known to exploit vulnerabilities in the same period.

While the number of malicious code samples that propagate by exploiting vulnerabilities has decreased this period, the number is still significant. Many of the samples exploiting vulnerabilities this period were bots. Bots can allow a remote attacker to perform numerous actions on a compromised computer, including stealing confidential information, launching DoS attacks, and installing additional threats.

Another growing shift in malicious code is in how it is reaching users. Traditionally, malicious code was delivered to the intended target. However, increasingly, malicious code samples are installed by attackers who lure users into visiting Web pages that exploit vulnerabilities in the user's browser or its components. The malicious code itself does not directly exploit any vulnerabilities in this scenario, but instead, is installed on a computer after a vulnerability is exploited.

For example, during the current reporting period, the MPack kit was used to install malicious code on computers.[146] Legitimate Web sites were compromised and legitimate Web pages were modified to include code to redirect the user's browser to a malicious server. The malicious MPack server then attempted to exploit one of a number of vulnerabilities to install the first stage of a multistaged downloader on the compromised computer.

This shift towards malicious code being installed through browser vulnerabilities can present challenges to network administrators. The variety of Web browsers and the number of components and plug-ins available for each can be daunting to keep track of and patch.[147] Antivirus software can detect malicious code samples that are installed by exploiting vulnerabilities. IPS technologies can also prevent exploitation of browser and plug-in vulnerabilities through signatures and behavior-based detection, as well as ASLR.

Percentage of computers with multiple infections

For the first time, in this volume of the *Internet Security Threat Report*, Symantec is assessing the number of times potential malicious code infections are reported from the same computer. This is done using data gathered by proprietary Symantec technologies. While many users may only experience one or two malicious code instances on their computers, some may become infected frequently within a single six-month period. Multiple infections may be due to a lack of knowledge on the user's part or out-of-date antivirus definitions. In some cases, multiple infections may also indicate that the computer was infected by a staged downloader, which will be discussed in the "Staged downloaders" metric below.

In the current period, 65 percent of computers reporting potential malicious code infections reported only a single instance of malicious code (figure 32). Thus, the majority of potentially infected users are likely to only experience a single malicious code instance in a period. This may be because many experienced computer users now make it a practice to update their antivirus signatures regularly. It should be noted that this data only takes into account malicious code infections over the current six-month period. A computer that only reported a single infection in the current period may have reported one in the previous period or may report one in the next.

[146] http://www.symantec.com/enterprise/security_response/weblog/2007/05/mpack_packed_full_of_badness.html
[147] For a more detailed discussion of Web browser plug-in vulnerabilities, see the "Vulnerability Trends" section of this report.

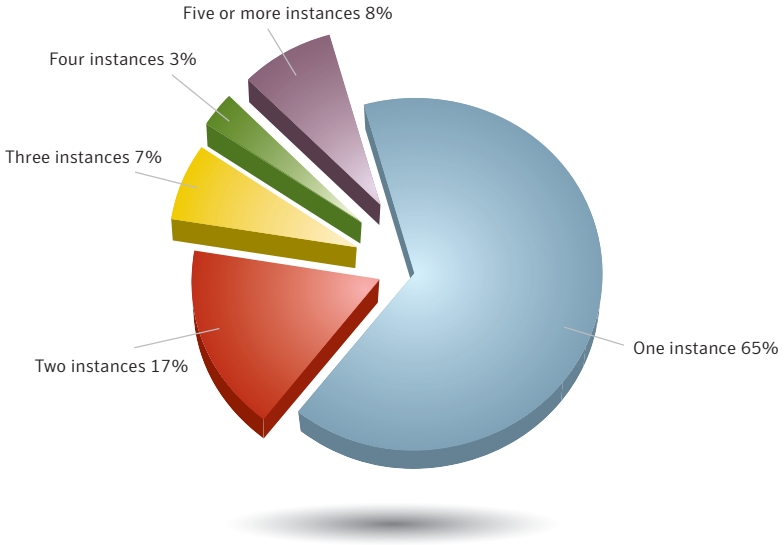

Figure 32. Percentage of computers with multiple infections
Source: Symantec Corporation

Thirty-five percent of computers reporting potential malicious code infections this period reported more than once. Seventeen percent of all computers reporting potential infections reported two potential infections. Some of these computers may report two potential infections because of staged downloaders or malicious code that downloads a second component. For example, the Mixor.Q worm also downloaded copies of the Peacomm Trojan on compromised computers.

Worth noting is that, in the first half of 2007, eight percent of computers reporting potential malicious code infections reported five or more potential infections. These users may engage in higher risk online behavior, such as following unknown links posted in forums, which could lead to malicious Web sites, or not keeping the patch levels of their software up-to-date. This type of behavior presents a risk to other users, particularly in corporate environments. A single compromised computer can potentially facilitate the infection of other users and servers on the network.

Users who experience multiple infections increase their likelihood of suffering serious consequences. Each time they are infected, they risk the theft of confidential information or loss of data. While a user may discover the first infection before the malicious code is able to send personal information back to the attacker, they might not be as fortunate with subsequent infections. For example, in the case of a staged downloader, the first infection may disable the security applications on a compromised computer, while the second infection contains a keystroke logger or some other remote access threat.

Staged downloaders—multiple infections by type

Staged downloaders, sometimes called modular malicious code, are threats that download and install other malicious code onto a compromised computer. These threats allow an attacker to change the downloadable component to any type of threat that suits their objectives. As the attacker's objectives change, he or she can change any later components that will be downloaded to perform the requisite tasks.

In the first half of 2007, the most prevalent downloader component was the Zlob Trojan (table 8).[148] This Trojan sets the user's Internet Explorer home, search, and "not found" pages to Web pages hosting malicious code. It also periodically displays fake security alerts from the System Tray that claim that the computer is infected. If the user clicks one of the error messages they will be directed to a Web page hosting malicious code.

Rank	Sample	Type	Download Mechanism
1	Zlob	Trojan	Redirects browser to malicious Web page
2	Vundo	Trojan	Downloads files from remote addresses
3	Mixor.Q	Worm	Downloads files from remote addresses
4	Anicmoo	Trojan	Downloads files from remote addresses
5	Skintrim	Trojan	Downloads files from remote addresses
6	Metajuan	Trojan	Downloads files from remote addresses
7	Stration	Worm	Downloads files from remote addresses
8	Wimad	Trojan	Uses Microsoft Windows Media® Digital Rights Manager to trick user into downloading files
9	Nebuler	Trojan	Downloads files from remote addresses
10	Secup	Trojan	Displays fake security alerts to trick user into downloading files

Table 8. Top staged downloaders
Source: Symantec Corporation

The Vundo Trojan was the second most common staged downloader by potential infections this period. Once Vundo is installed on a computer, it attempts to contact certain IP addresses to download and install its secondary components. One of the files it attempts to install is an adware program that will cause pop-up advertisements to be displayed periodically. The adware component likely provides revenue to the malicious code author.

Mixor.Q was the third most common staged downloader in the first six months of 2007. It is a mass-mailing worm that was also responsible for part of the Peacomm outbreak. This worm sends out a mass-mailing of itself in order to propagate and is also known to install either Peacomm or Galapoper.A.[149] Both of these secondary stages also download additional threats onto compromised computers. Galapoper can also be used to relay spam.

All of the top ten second-stage components downloaded this period were Trojans or back doors. Some of these Trojans simply download another threat to the compromised computer, while others steal confidential information or compromise the computer's security, leaving it open to further compromise.

[148] http://www.symantec.com/security_response/writeup.jsp?docid=2005-042316-2917-99
[149] http://www.symantec.com/security_response/writeup.jsp?docid=2006-042013-1813-99

The most prevalent downloaded component in the first six months of 2007 was the Adclicker Trojan (table 9).[150] This simple Trojan is intended to drive traffic to Web pages and banner advertisements. Banner advertisements compensate the owner of the Web site they are hosted on for each view or click-through.[151] Generating fraudulent traffic to these advertisements is commonly referred to as click fraud.

Rank	Sample	Type	Impact
1	Adclicker	Trojan	Generates traffic to Web sites and banner ads
2	Gampass	Trojan	Steals online gaming account information
3	Zonebac	Trojan	Lowers Internet Explorer security settings
4	KillAV	Trojan	Disables security applications
5	Lineage	Trojan	Steals online gaming account information
6	Peacomm	Trojan	Creates a peer network and downloads other threats
7	Rustock.B	Back door	Allows remote access and relays spam
8	Bzup	Trojan	Steals online banking account information
9	Graybird	Back door	Allows remote access, logs keystrokes, and steals passwords
10	Haxdoor	Back door	Allows remote access

Table 9. Top downloaded components
Source: Symantec Corporation

The Gampass Trojan was the second most commonly downloaded component this period.[152] It is primarily used to steal a user's online gaming account information and send it to the attacker. This Trojan is discussed in greater detail in the "Malicious code targeting online gaming" section below.

Zonebac was the third most commonly downloaded component in the first six months of 2007. It is a Trojan that lowers the Internet Explorer security zone settings.[153] These settings prevent Web sites from automatically downloading and executing files through the browser. Zonebac also starts a hidden process to connect to certain Web sites, which will likely attempt to take advantage of the lowered security zone settings to install other threats on the compromised computer.

All of the top ten staged downloaders and eight of the top ten downloaded components were also among the top 50 malicious code samples by potential infections this period. Twenty-eight of the top 50 samples accounting for 79 percent of potential infections included the ability to download additional components. This illustrates the prevalence of staged downloaders during the current reporting period. Since many staged downloaders consist of Trojans, this also relates to the increase in Trojans causing potential infections as discussed in the "Malicious code types" section of this report.

[150] http://www.symantec.com/security_response/writeup.jsp?docid=2002-091214-5754-99
[151] A click-through is a link that contains uniquely identifiable information about its originator that a user clicks on. Typically, the originator receives financial compensation for each click-through.
[152] http://www.symantec.com/security_response/writeup.jsp?docid=2006-111201-3853-99
[153] http://www.symantec.com/security_response/writeup.jsp?docid=2006-091612-5500-99

Malicious code targeting online gaming

Online gaming is becoming one of the most popular activities on the Internet. Recently, a study indicated that unique visitors to online gaming sites reached 217 million worldwide.[154] In 2007, the online gaming market in China, where there were 30 million Internet gamers by the end of 2006,[155] is expected to grow by 35 percent.[156]

Online games often feature goods, such as prizes, that can be exchanged by players for money. The total annual wealth created within virtual worlds has been placed at approximately 10 billion USD.[157] As such, it is not surprising that attackers appear to be turning their attention to these games. This metric will assess malicious code that targets online gaming, including:

- The top three malicious code samples targeting online gaming sites
- The percentage of the top 50 malicious code samples that target these sites
- The most commonly targeted gaming sites

In the first half of 2007, the most common malicious code sample targeting online games was the Gampass Trojan (table 10). This Trojan is notable because the attacker can use it to target one of several online games, including the Lineage, Ragnarok Online, Rohan, and Rexue Jianghue games. These games are more popular in the APJ region than the rest of the world. As a result, 84 percent of worldwide potential infections by Gampass during this period originated in that region.

The ability of this threat to be configured to target multiple games likely contributes to its popularity among attackers. When it is installed, the Trojan will log keystrokes when the user connects to a specified online gaming site. It will then send the log to a Web site or email address. Gampass may also attempt to disable the processes of antivirus and other security products, leaving compromised users open to additional threats.

Sample	Type	Game(s) Targeted
Gampass	Trojan	Configurable for many
Lineage	Trojan	Lineage
Dowiex	Virus, Trojan	World of Warcraft

Table 10. Top three malicious code samples targeting online gaming sites
Source: Symantec Corporation

The second most common malicious code sample targeting online games this period was the Lineage Trojan.[158] This Trojan steals account information for the Lineage online game and emails it to the attacker. Interestingly, this Trojan was first seen on January 11, 2005, yet it still remains one of the top 50 malicious code samples reported to Symantec two years later. The persistence of this Trojan is likely due to the fact that authors are continually creating new variants to bypass antivirus signatures. This indicates that the Trojan has been proven to be effective and successful, or attackers would most likely have created newer threats to accomplish their goals.

[154] http://www.comscore.com/press/release.asp?press=1521
[155] http://abcnews.go.com/Technology/wireStory?id=3386396
[156] http://uk.reuters.com/article/internetNews/idUKSHA27160820070628
[157] http://www.pcworld.com/article/id,128270-page,2-c,onlineentertainment/article.html
[158] http://www.symantec.com/security_response/writeup.jsp?docid=2005-011211-3355-99

Dowiex was the third most common malicious code sample targeting online games during the first six months of 2007.[159] This threat downloads the Wowcraft Trojan[160] onto compromised computers. This Trojan, in turn, logs keystrokes in windows with certain titles associated with the World of Warcraft game. Like Gampass, Wowcraft also disables processes associated with security applications. It can also download and install other threats on the compromised computer.

In the first six months of 2007, five percent of the top 50 malicious code samples reported to Symantec attempted to steal account information for online games. This is likely due to the fact that there is considerable financial gain to be made from online gaming accounts, so that attackers are deploying these threats in substantial numbers.

Another indication of the growing appeal of targeting online gaming is that both Gampass and Lineage were also two of the most downloaded components of multistaged downloaders this period. This indicates that attackers see value in targeting online gamers since many of the other top downloaded components are used for more common types of identity theft such as stealing online banking account credentials.

Of further concern is that two of the top three malicious code threats targeting online games also disable security applications on the compromised computer. This could leave the computer open to other threats even if the user does not participate in any of these online games. Combined with the ability to download other threats, this means that attackers can install a wide range of threats on compromised computers once they have the user's online gaming account information.

Malicious code—prevention and mitigation

Symantec recommends that certain best security practices always be followed to protect against malicious code infection. Administrators should keep patch levels up to date, especially on computers that host public services and applications—such as HTTP, FTP, SMTP, and DNS servers—and that are accessible through a firewall or placed in a DMZ. Email servers should be configured to only allow file attachment types that are required for business needs and to block email that appears to come from within the company, but that actually originates from external sources. Additionally, Symantec recommends that ingress and egress filtering be put in place on perimeter devices to prevent unwanted activity.

To protect against malicious code that installs itself through a Web browser, additional measures should be taken. The use of IPS technologies can prevent exploitation of browser and plug-in vulnerabilities through signatures and behavior-based detection in addition to ASLR.

End users should employ defense-in-depth strategies, including the deployment of antivirus software and a personal firewall. Users should update antivirus definitions regularly. They should also ensure that all desktop, laptop, and server computers are updated with all necessary security patches from their software vendors. They should never view, open, or execute any email attachment unless it is expected and comes from a trusted source, and unless the purpose of the attachment is known.

[159] http://www.symantec.com/security_response/writeup.jsp?docid=2006-101716-2136-99
[160] http://www.symantec.com/security_response/writeup.jsp?docid=2005-073115-1710-99

Phishing Trends

Phishing is an attempt by a third party to solicit confidential information from an individual, group, or organization by spoofing a specific, well known brand, usually for financial gain. Phishers are groups or individuals who attempt to trick users into disclosing personal data, such as credit card numbers, online banking credentials, and other sensitive information. They may then use the information to commit fraudulent acts. This section of the *Internet Security Threat Report* will discuss phishing activity that Symantec detected between January 1 and June 30, 2007.

The data provided in this section is based on statistics derived from the Symantec Probe Network, which consists of over two million decoy email accounts that attract email messages from 20 different countries around the world. The main purpose of the network is to attract spam, phishing, viruses, and other email-borne threats. It encompasses more than 600 participating enterprises around the world, attracting email that is representative of traffic that would be received by over 250 million mailboxes. The Probe Network consists of previously used email addresses as well as email accounts that have been generated solely to be used as probes.

In addition to the Probe Network, Symantec also gathers phishing information through the Symantec Phish Report Network, an extensive antiphishing community of enterprises and consumers.[161] Members of the Phish Report Network contribute and receive fraudulent Web site addresses for alerting and filtering across a broad range of solutions.

Symantec assesses phishing according to two indicators: phishing messages and phishing attempts. A phishing message is a single, unique message that is sent to targets with the intent of gaining confidential and/or personal information from computer users by directing them to a Web site where the user's information is fraudulently obtained. Each phishing message has different content and each one will represent a different way of trying to fool a user into disclosing information by spoofing a known brand. A phishing message can be considered the "lure" with which a phisher attempts to entice a phishing target to disclose confidential information.

A phishing attempt can be defined as an instance of a phishing message being sent to a single user. Extending the fishing analogy, a phishing attempt can be considered a single cast of the lure (the phishing message) to try to ensnare a target. A single phishing message can be used in numerous distinct phishing attempts, usually targeting different end users.

[161] http://www.phishreport.net

Phishing Highlights

The following section will offer a brief summary of some of the phishing trends that Symantec observed during this period, based on data provided by the sources listed above. Following this overview, the *Internet Security Threat Report* will discuss selected phishing metrics in greater depth, providing analysis and discussion of the trends indicated by the data.

- The Symantec Probe Network detected a total of 196,860 unique phishing messages, an 18 percent increase over the last six months of 2006. This equates to an average of 1,088 unique phishing messages per day for the first half of 2007.
- Symantec blocked over 2.3 billion phishing messages in the first half of 2007, an increase of 53 percent over the second half of 2006. This means that Symantec blocked an average of roughly 12.5 million phishing emails per day over the first six months of 2007.
- Organizations in the financial services sector accounted for 79 percent of the unique brands that were used in phishing attacks during this period.
- The brands of organizations in the financial services sector were spoofed by 72 percent of all phishing Web sites.
- Fifty-nine percent of all phishing Web sites detected in the first half of 2007 were located in the United States, a much higher proportion than in any other country.
- Three phishing toolkits were responsible for 42 percent of all phishing attacks observed by Symantec in the first half of 2007.
- Eighty-six percent of all known phishing Web sites were hosted on only 30 percent of IP addresses known to be phishing Web servers.

Phishing Discussion

This section will discuss selected phishing metrics in greater depth, providing analysis and discussion of the trends indicated by the data. The following metrics will be discussed:

- Phishing activity by sector
- Top countries hosting phishing Web sites
- Automated phishing toolkits
- Core brands being phished
- Phishing—prevention and mitigation

Phishing activity by sector

This metric will assess phishing activity by sector. It will do this in two ways. First, it will identify the sectors in which the organizations that were most commonly phished belong. This means that the organization's brand was used in phishing attacks. Second, it will assess which sectors were targeted by the highest volume of phishing attacks. These considerations are important for enterprises because the use of an organization's brand in phishing activity can have significant negative consequences. It can undermine consumer confidence and damage the organization's reputation. Furthermore, the company may be required to compensate victims of any phishing scams that use the company's brand.

Most of the organizations whose brands were used in phishing attacks in the first six months of 2007 were part of the financial services sector. Organizations in that sector accounted for 79 percent of the brands that were used for phishing during this period (figure 33), compared to the previous period when they accounted for 84 percent. The financial services sector also accounted for the highest volume of phishing Web sites during this period, making up 72 percent of all phishing Web sites reported to Symantec (figure 34). Financial services made up 64 percent of all phishing Web sites in the last half of 2006.

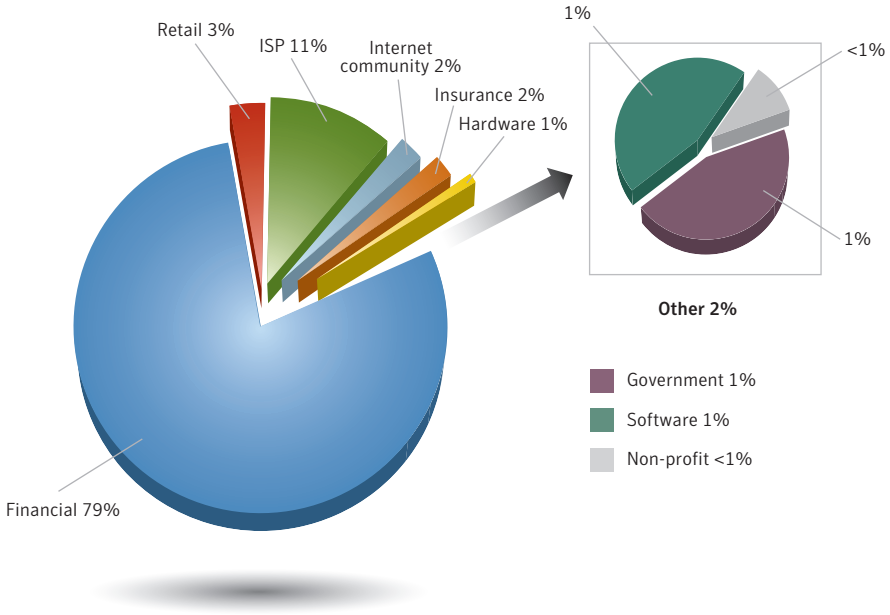

Figure 33. Brands phished by sector
Source: Symantec Corporation

Most phishing activity is conducted for financial gain. A successful phishing attack that mimics the brand of a financial entity is most likely to yield data that can be used for immediate financial gain. It is therefore logical that phishing attacks focus on brands within the financial services sector.

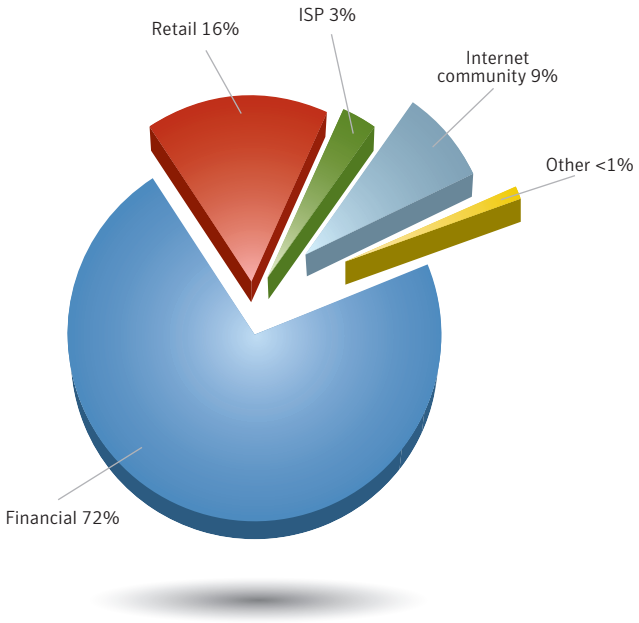

Figure 34. Phished sectors by phishing Web sites
Source: Symantec Corporation

Organizations in the Internet service provider (ISP) sector made up 11 percent of the unique brands used in phishing attacks during this period, making it the second ranked sector. This is an increase over the seven percent of phishing attacks that spoofed ISP brands in the second half of 2006.

As noted in the previous edition of the *Internet Security Threat Report*, ISP accounts can be valuable targets for phishers.[162] People frequently use the same authentication credentials (such as usernames and passwords) for multiple accounts, including their email accounts.[163] Thus, information gleaned through phishing attacks may provide access to other accounts, such as online banking.

Additionally, attackers could use the free Web-hosting space that is often provided with these accounts to host phishing Web sites, or they could use the accompanying email accounts to send spam or launch further phishing attacks. In some cases, compromised ISP Web-hosting may also be used to plant links to other Web sites the attacker controls in order to boost the rating of the Web site in search engines.[164] Email account passwords were also the third most common item advertised for sale on underground economy servers this period, as described in the "Underground economy servers" discussion in the "Attack Trends" section of this report.

The retail services sector only accounted for three percent of organizations whose brands were spoofed in phishing activity in the first half of 2007; however, it accounted for 16 percent of the volume of phishing Web sites. In the previous reporting period, it accounted for five percent of the unique brands spoofed and 34 percent of phishing Web sites.

[162] Symantec *Internet Security Threat Report*, Volume XI (March 2007): http://eval.symantec.com/mktginfo/enterprise/white_papers/ent-whitepaper_internet_security_threat_report_xi_03_2007.en-us.pdf : p. 69
[163] http://cups.cs.cmu.edu/soups/2006/proceedings/p44_gaw.pdf
[164] For a more detailed discussion of search engine rankings, see the "Malicious Code Trends" section of this report.

The disproportionate number of phishing Web sites in the retail services sector indicates that a small number of retail brands were being heavily phished. This is illustrated by the fact that a large volume of phishing attacks were reported that attempted to spoof the eBay brand. This is not surprising, as an attacker can use a user's eBay account credentials in various ways. First, many eBay accounts are linked to the user's PayPal account. As users often use the same passwords for these accounts, compromising one could give an attacker access to both, which would allow the attacker to transfer funds to him- or herself. Additionally, the attacker could use the account to buy goods from other users and default on the transaction, sell items that do not exist, or even use the account to sell stolen goods or goods purchased from an online retailer using a hijacked account or stolen credit card.

While the retail services sector made up 16 percent of phishing Web sites, this is a significant decrease from the 34 percent reported in the previous six-month period. This is mainly due to a significant rise in the volume of phishing sites targeting the financial sector. Attackers have also started exploring other means of perpetrating fraud upon customers of retail organizations such as eBay. For instance, some Trojans and other attacks[165] can also facilitate identity theft.

Eight of the top ten brands spoofed by attackers in phishing attacks during this period were in the financial sector. Interestingly, one of the most frequently spoofed brands this period was an Internet community. While there is no immediate financial gain to be obtained by attackers who steal a user's account information, it may provide other returns. The attacker could use the account to gather information from the hijacked account's friends, such as email addresses, by sending messages that appear to come from the legitimate user, who would likely be implicitly trusted by the message recipient.[166] Additionally, the attacker can send messages containing links to Web sites that are designed to download malicious code on visitors' computers.[167] Since the link comes from a user's friend, they may be more likely to trust the link and visit the site.[168]

Top countries hosting phishing Web sites

A phishing Web site is a site that is designed to mimic the legitimate Web site of the organization whose brand is being spoofed, often an online bank or e-commerce retailer. In many cases, it is set up by the attacker to capture a victim's authentication information or other personal identification information, which can subsequently be used in identity theft or other fraudulent activity.

This metric will assess the countries in which the most phishing Web sites were hosted in the first six months of 2007. In this case, Symantec counts phishing Web sites as the number of unique IP addresses hosting Web pages used for phishing. This data is a snapshot in time, and does not offer insight into changes in the locations of certain phishing sites over the course of the reporting period. It should also be noted that the fact that a phishing Web site is hosted in a certain country does not necessarily mean that the attacker is located in that country.

[165] Please see http://www.symantec.com/enterprise/security_response/weblog/2007/03/ebay_motor_scam_update.html and http://redtape.msnbc.com/2007/03/how_far_has_vla.html, respectively, for more in-depth discussions.
[166] http://www.symantec.com/enterprise/security_response/weblog/2006/11/an_imaginative_phishing_attack_1.html
[167] http://blog.washingtonpost.com/securityfix/2007/06/web_2pointuhoh_worm_whacks_mys.html
[168] For more on phishing attacks that target social networking sites, please see: http://www.symantec.com/enterprise/security_response/weblog/2006/09/contextaware_phishing_realized.html

In the first half of 2007, 59 percent of all known phishing Web sites were located in the United States (table 11), a considerable increase over the previous period when 46 percent of phishing Web sites were located there. The United States is home to a large number of Web-hosting providers, including over 30 percent of registered domains.[169] It is also home to the highest number of Internet users in the world.[170] The increase in phishing Web sites located there during this reporting period is also likely related to the high number of Trojans reported from North America this period, as is discussed in the "Malicious Code Trends" section of this report. Trojans are frequently used for hosting Web sites used in phishing attacks.

Rank	Previous Rank	Country	Current Period	Previous Period
1	1	United States	59%	46%
2	2	Germany	6%	11%
3	3	United Kingdom	3%	3%
4	10	Netherlands	2%	2%
5	11	Russia	2%	2%
6	4	France	2%	3%
7	7	Canada	2%	2%
8	5	Japan	2%	3%
9	8	China	1%	2%
10	6	Taiwan	1%	3%

Table 11. Top countries hosting phishing Web sites
Source: Symantec Corporation

Germany was once again the location of the second-highest percentage of phishing Web sites this period, with six percent of the worldwide total. This is, however, a decrease from the last six months of 2006 when 11 percent of phishing Web sites were located there. Variations in percentages between periods are likely a result of the opportunistic nature of attackers. Attackers are most likely to host phishing Web sites on any computer they are able to compromise. In many cases, attackers host their phishing Web sites on a computer that was compromised by a bot. Because bots compromise any computer that is vulnerable to the exploits they use to propagate, there is little control on the part of the attacker as to the physical location of computers in their bot network.

The United Kingdom hosted the third highest number of phishing Web sites this period. It held steady at three percent of worldwide phishing Web sites reported in the previous period. The percentage of bots in the United Kingdom has been dropping in recent periods; however, it is the top country reporting potential malicious code infections in the EMEA region. This may indicate that attackers are using bots less frequently in phishing attacks and are instead using other malicious code to host phishing Web sites.

[169] http://www.webhosting.info/webhosts/tophosts/global/
[170] http://www.pewinternet.org/PPF/r/218/report_display.asp

Automated phishing toolkits

For the first time, in this volume of the *Internet Security Threat Report*, Symantec is assessing the usage of automated phishing toolkits. A phishing toolkit is a set of scripts that allows an attacker to automatically set up phishing Web sites that spoof the legitimate Web sites of different brands, including the images and logos associated with those brands. The scripts also help to generate corresponding phishing email messages. As each script generates pseudo-random phishing URLs with a distinctive pattern, the particular script used to generate a particular phishing URL can be identified from that pattern. All phishing URLs reported to Symantec can be sorted and grouped according to those specific patterns.

Phishing toolkits are developed by groups or individuals and are sold on the underground market. As such, they illustrate the trend that Symantec has observed towards an increase in the commercialization, development and distribution in threats and malicious services. This trend also indicates that phishing is becoming an increasingly organized activity. These sophisticated phishing kits are sold for a lot of money, so it's unlikely they would be available to an average user.

The three phishing kits examined in this discussion are quite a bit more robust than others Symantec has analyzed. For example, these kits include tools to construct the phishing Web sites and they allow multiple phishing Web sites to be created on the same compromised computer. They also enable the attacker to automate the creation and sending of the phishing email messages. Other kits often only include scripts to send email messages or tools for creating the phishing Web site.

A look at the three most widely used phishing toolkits reveals that, on average, they alone were responsible for 42 percent of all phishing attacks detected in the first half of 2007 (figure 35).[171] This shows the high percentage of complete automation used in phishing attacks compared to attacks that are only partially automated. Automation allows attackers to send a high volume of phishing messages that spoof several brands to a large number of recipients with minimal effort. Of the 58 percent of remaining attacks, some may have used phishing toolkits other than the three that are currently known to Symantec, while others used techniques other than toolkits.

[171] It should be noted that most of the remaining phishing attacks likely use simple scripts at some point in their attack process to simplify certain repetitive tasks, but for this analysis, the focus was on the three most widely used and completely automated phishing toolkits that generate pseudo-random URL links.

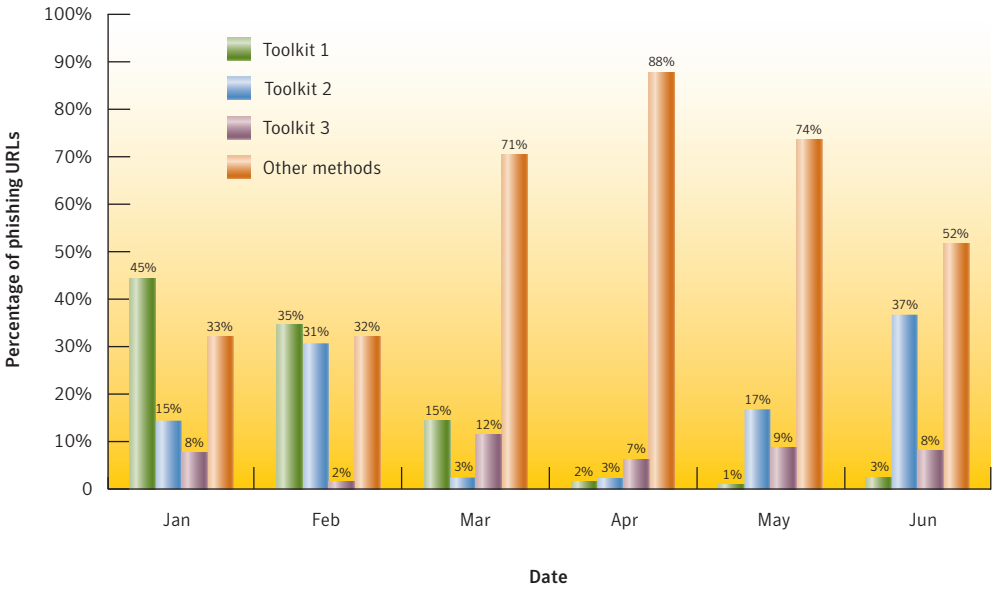

Figure 35. Use of automated phishing toolkits
Source: Symantec Corporation

There is significant dynamism in the toolkits that are used at any one point in time. For instance, Toolkit 1 declined from 45 percent usage in January to only three percent usage in June. This suggests that the attackers using it may have stopped because they moved to a different version of this toolkit or some other entirely different toolkit. This adaptation is typical as new detection and protection methods are introduced over time.

One indicator that a phishing toolkit has been used is that a number of phishing Web sites are hosted on a single IP address. A toolkit can easily set up phishing Web sites that spoof a number of different brands on the same compromised computer. Hosting multiple phishing Web sites on a single computer offers numerous advantages. For instance, the attacker doesn't need to worry about maintaining multiple computers and can instead use a toolkit to easily host Web sites that mimic several brands on the single computer. However, doing so creates a single point of failure for the attacker. If authorities discover the host computer before the attacker can gather the information collected from victims, he or she loses much more data than if each phishing site had been hosted on a separate computer.

During the first half of 2006, 86 percent of all phishing Web sites reported to Symantec were hosted on only 30 percent of phishing IP addresses. Examining the data throughout the period reveals a strong link between the number of phishing IP addresses and the use of phishing toolkits, as described above. There is a strong correlation between months with a high number of phishing URLs not generated by toolkits and months with a decrease in phishing Web sites hosted on the same IP addresses (figure 36).

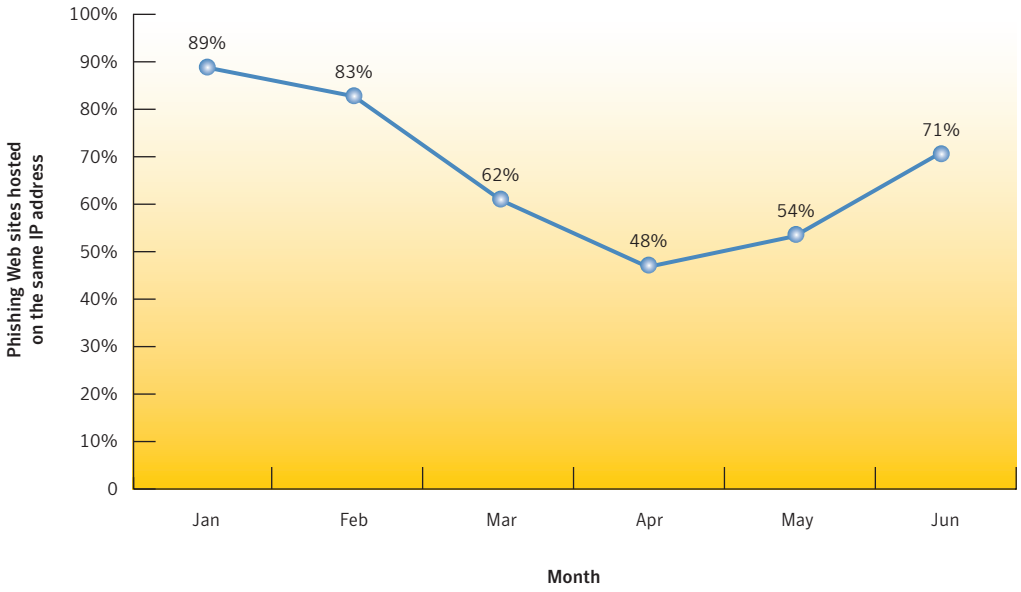

Figure 36. Phishing Web sites that use the same IP address
Source: Symantec Corporation

The percentage of phishing Web sites hosted on a single IP address was high during January and February, but was significantly lower from March through May. These months also saw a decrease in phishing URLs that were generated by the three major phishing kits and an increase in phishing URLs generated by other means. For example, in April only 48 percent of phishing Web sites were hosted on the same IP address. This is consistent with the use of phishing toolkits to create phishing Web sites for multiple brands on a single computer.

In addition to phishing toolkits, the use of Web hosting services to host phishing Web sites also contributes to multiple phishing Web sites residing on the same IP address. While one or more attackers may use multiple accounts with the same hosting company to host phishing Web sites, they may still physically reside on the same server or on a group of servers using the same gateway IP address. This can also present difficulties for the Web hosting provider. Since their IP addresses can potentially be included on DNS block lists if a phishing Web site resides on their servers, this can also cause legitimate Web sites they host to be blocked.

Core brands being phished

For the first time, in this volume of the *Internet Security Threat Report*, Symantec is analyzing the core brands that were spoofed, or mimicked, during phishing attacks. Core brands are brands that are spoofed at least once each month in a phishing attack. These core brands were determined by identifying six lists of brands, one per month from January through June 2007, in which a new Web site spoofing that brand was reported. The core brands, then, are those that were present on each of these lists. In other words, the core brands were those for which a new phishing Web site was known to have been created in each month of this reporting period.

During the first six months of 2007, Symantec classified 78 of the 359 brands mimicked in a phishing attack as core brands. Symantec then compared the core brands with the most frequently spoofed brands; that is, the brands for which the greatest number of spoofed phishing Web sites were detected. While many core brands are among the most frequently spoofed brands, there are also significant differences.

In particular, among the top 78 most frequently spoofed brands, only 61 were core brands. The ninth most frequently spoofed brand was actually not a core brand. This is primarily because Symantec did not see any reports of Web sites spoofing this specific brand during the month of April. The eleventh most frequently spoofed brand was also not a core brand. In fact, this brand was only spoofed during the month of February, and 98 percent of the phishing sites that spoofed it were observed during a one-week period. This spike in activity might suggest that phishers unearthed some temporary security weakness in the site, such as an easy cash-out mechanism, and decided to target it. The weakness might have been shored up or the phishers efforts might have otherwise been unsuccessful, causing them to look elsewhere.

At the other end of the spectrum, the least frequently spoofed core brand was ranked 112th out of 359 among the most frequently spoofed brands. Only 12 phishing sites were set up to spoof this core brand. Three phishing sites spoofing that brand were reported in each of January, March, and April, and only one new site was reported in each of February, May, and June.

These numbers suggest that phishers do not always take a scatter-shot approach in their attack attempts. Instead, for specific targets, they prefer methodical smaller-scaled approaches, albeit at a consistent pace. In general, the data seems to suggest that phishers vary their approach depending on the brand. Some brands are continuously spoofed whereas others are consistently, but less frequently, spoofed. Phishers may be adapting their behavior to optimize for profitability.

Phishing—prevention and mitigation

Symantec recommends that enterprise users protect themselves against phishing threats by filtering email at the server level through the mail transfer agent (MTA). Although this will likely remain the primary point of filtering for phishing, organizations can also use IP-based filtering upstream, as well as HTTP filtering.

DNS block lists also offer protection against potential phishing emails.[172] Organizations could also consider using domain-level or email authentication in order to verify the actual origin of an email message. This can protect against phishers who are spoofing email domains.[173]

To protect against potential phishing activity, administrators should always follow Symantec best practices as outlined in Appendix A of this report. Symantec also recommends that organizations educate their end users about phishing.[174] They should also keep their employees notified of the latest phishing attacks and how to avoid falling victim to them, as well as provide a means to report suspected phishing sites.[175]

Organizations can also employ Web server log monitoring to track if and when complete downloads of their Web sites, logos, and images are occurring. Such activity may indicate that someone is using the legitimate Web site to create an illegitimate Web site that could be used for phishing.

[172] A DNS block list (sometimes referred to as a black list) is simply a list of IP addresses that are known to send unwanted email traffic. It is used by email software to either allow or reject email coming from IP addresses on the list.
[173] Spoofing refers to instances where phishers forge the "From:" line of an email message using the domain of the entity they are targeting with the phishing attempt.
[174] For instance the United States Federal Trade Commission has published some basic guidelines on how to avoid phishing. They are available at: http://www.ftc.gov/bcp/edu/pubs/consumer/alerts/alt127.htm
[175] A good resource for information on the latest phishing threats can be found at: http://www.antiphishing.org

Organizations can detect phishing attacks that use spoofing by monitoring non-deliverable email addresses or bounced email that is returned to non-existent users. They should also monitor the purchasing of cousin domain names by other entities to identify purchases that could be used to spoof their corporate domains.[176] So-called typo domains[177] and homographic domains[178] should also be monitored. This can be done with the help of companies that specialize in domain monitoring; some registrars also provide this service.

The use of anti-phishing toolbars and components in Web browsers can also help protect users from phishing attacks. These measures notify the user if a Web page being visited does not appear to be legitimate. This way, even if a phishing email reaches a user's inbox, the user can still be alerted to the potential threat.

End users should follow best security practices, as outlined in Appendix A of this report. They should use an antiphishing solution. As some phishing attacks may use spyware and/or keystroke loggers, Symantec advises end users to use antivirus software, antispam software, firewalls, toolbar blockers, and other software detection methods. Symantec also advises end users to never disclose any confidential personal or financial information unless and until they can confirm that any request for such information is legitimate.

Users should review bank, credit card, and credit information frequently. This can provide information on any irregular activities. For further information, the Internet Fraud Complaint Center (IFCC) has also released a set of guidelines on how to avoid Internet-related scams.[179] Additionally, network administrators can review Web proxy logs to determine if any users have visited known phishing sites.

[176] Cousin domains refers to domain names that include some of the key words of an organization's domain or brand name; for example, for the corporate domain "bigbank.com", cousin domains could include "bigbank-alerts.com", "big-bank-security.com", and so on.
[177] Typo domains are domain names that use common misspellings of a legitimate domain name, for example the domain "symatnec.com" would be a typo domain for "symantec.com".
[178] A homographic domain name uses numbers that look similar to letters in the domain name, for example the character for the number "1" can look like the letter "l".
[179] http://www.fbi.gov/majcases/fraud/internetschemes.htm

Spam Trends

Spam is usually defined as junk or unsolicited email sent by a third party. While it is certainly an annoyance to users and administrators, spam is also a serious security concern as it can be used to deliver Trojans, viruses, and phishing attempts.[180] It could also cause a loss of service or degradation in the performance of network resources and email gateways. This section of the *Internet Security Threat Report* will discuss developments in spam activity between January 1 and June 30, 2007.

The data used in this analysis is based on data returned from the Symantec Probe Network as well as data gathered from a statistical sampling of the Symantec Brightmail AntiSpam customer base. Specifically, statistics are gathered from enterprise customers' Symantec Brightmail AntiSpam servers that receive more than 1,000 email messages per day. This removes the smaller data samples (that is, smaller customers and test servers), thereby allowing for a more accurate representation of data.

The Symantec Probe Network consists of millions of decoy email addresses that are configured to attract a large stream of spam attacks. An attack can consist of one or more messages. The goal of the Probe Network is to simulate a wide variety of Internet email users, thereby attracting a stream of traffic that is representative of spam activity across the Internet as a whole. For this reason, the Probe Network is continuously optimized in order to attract new varieties of spam attacks. This is accomplished through internal production changes that are made to the network, which thus affect the number of new spam attacks it receives as a whole.

Spam Highlights

The following section will offer a brief summary of some of the spam trends that Symantec observed during this period based on data provided by the sources listed above. Following this overview, this section will discuss selected metrics in greater depth, providing analysis and discussion of the trends indicated by the data.

- Between January 1 and June 30, 2007, spam made up 61 percent of all email traffic monitored at the gateway. This is a slight increase over the last six months of 2006 when 59 percent of email was classified as spam.
- Sixty percent of all spam detected during this period was composed in English, down from 65 percent in the previous reporting period.
- In the first half of 2007, 0.43 percent of all spam email contained malicious code, compared to 0.68 percent of spam that contained malicious code in the second half of 2006. This means that one out of every 233 spam messages blocked by Symantec Brightmail AntiSpam contained malicious code.
- Spam related to commercial products made up 22 percent of all spam during this period, the most of any category.
- During the first six months of 2007, 47 percent of all spam detected worldwide originated in the United States compared to 44 percent in the previous period.
- The United States hosted the most spam zombies of any country, with 10 percent of the worldwide total.
- In the first half of 2007, 27 percent of all spam blocked by Symantec was image spam.

[180] http://news.bbc.co.uk/2/hi/technology/6676819.stm

Spam Discussion

This section will discuss selected spam metrics in greater depth, providing analysis and discussion of the trends indicated by the data. The following metrics will be discussed:

- Top spam categories
- Top countries of spam origin
- Image spam

Top spam categories

Spam categories are assigned by Symantec Email Security Group analysts based on spam activity that is detected by the Symantec Probe Network. While some of the categories may overlap, this data provides a general overview of the types of spam that are most commonly seen on the Internet today. It is important to note that this data is restricted to spam attacks that are detected and processed by the Symantec Probe Network. Internal upstream processing may weed out particular spam attacks, such as those that are determined to be potential fraud attacks.

The most common type of spam detected in the first half of 2007 was related to commercial products (figure 37), which made up 22 percent of all spam detected by Symantec sensors during this period. This is a slight increase from the previous period when this category made up 21 percent of detected spam. Commercial product spam usually consists of advertisements for commercial goods and services. It is frequently used to sell designer goods, such as watches, handbags, and sunglasses. There is financial motivation since the goods sold are often counterfeit and can be sold at a profit. In some cases the spammers may simply be collecting credit card and personal information for use in identity theft.

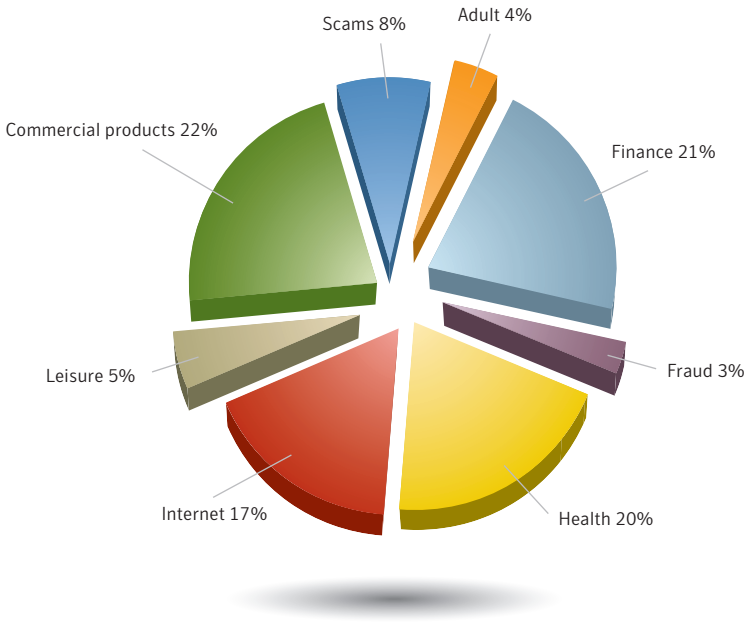

Figure 37. Top spam categories
Source: Symantec Corporation

Spam related to financial services made up 21 percent of all spam in the first six months of 2007, making it the second most common type of spam during this period. The previous edition of the *Internet Security Threat Report* reported that Symantec had detected an increase in spam related to the financial services sector over the last six months of 2006. This was primarily due to an abundance of stock market "pump and dump" spam.[181] However, in the current period, there has been a 30 percent decline in this type of spam from the previous period. This is due to a decline in spam touting penny stocks that was triggered by actions taken by the United States Securities and Exchange Commission,[182] which limited the profitability of this type of spam by suspending trading of the stocks that are touted.

Spam related to health products and services made up 20 percent of all spam detected during this period compared to 23 percent in the second half of 2006. This category traditionally has one of the highest click-through rates, as it tends to be more difficult to market through more legitimate and traditional means. A click-through is a link that is embedded in a spam message. The link contains uniquely identifiable information about its originator. Each time a user clicks on the link, it is considered a click-through. Typically, the originator receives financial compensation for each click-through. Spammers have an economic incentive to have a high click-through rate in order to increase their return on investment. Therefore, it is reasonable to conclude that they would use spam content that has a high click-through rate.

Internet-related spam rose to 17 percent this period from 10 percent in the last half of 2006. This type of spam is typically used to promote Web hosting and design, as well as other online commodities like phishing and spam toolkits. Since phishing and spam toolkits cannot typically be advertised by legitimate means, such as through banner ads on Web sites, spam tends to be the only way to promote them.

Top countries of spam origin

This section will discuss the top ten countries of spam origin. The nature of spam and its distribution on the Internet presents challenges in identifying the location of people who are sending it. Many spammers try to redirect attention away from their actual geographic location. In an attempt to bypass DNS block lists, they use Trojans that relay email, which allow them to send spam from sites that are distinct from their physical location. In doing so, they will likely focus on compromised computers in those regions with the largest bandwidth capabilities. Following this logic, the region in which the spam originates may not correspond with the region in which the spammers are located.

This discussion is based on data gathered by customer installations of Symantec Brightmail AntiSpam. This data includes the originating server's IP address, against which frequency statistics are summarized. Each IP address is mapped to a specific country and charted over time.

During the first six months of 2007, 47 percent of all spam detected worldwide originated in the United States (table 12). This is likely due to the high number of broadband users in that country and the high percentage of bot-infected computers located there, as was discussed in the "Attack Trends" section of this report. The United States was also the top country of spam origin in the second half of 2006, when 44 percent of spam originated there.

[181] For a more in-depth discussion of pump-and-dump spam, please see the Symantec *Internet Security Threat Report*, Volume 11 (March 2007): http://eval.symantec.com/mktginfo/enterprise/white_papers/ent-whitepaper_internet_security_threat_report_xi_03_2007.en-us.pdf : p.16
[182] http://www.sec.gov/news/press/2007/2007-34.htm

Top Ten Countries of Spam Origin	Jan–Jun 2007	Jul–Dec 2006
United States	47%	44%
Undetermined EU Countries	7%	7%
China	4%	6%
United Kingdom	4%	3%
Japan	4%	3%
South Korea	3%	3%
Taiwan	3%	1%
Poland	3%	3%
Germany	2%	2%
Switzerland	2%	1%

Table 12. Top ten countries of spam origin
Source: Symantec Corporation

The second highest source of spam this period was a group of undetermined European Union countries. Seven percent of all detected spam originated there this period, the same amount as the second half of 2006. In this group, the specific source countries cannot be definitively identified because the ISPs through whose networks the spam was sent operate in more than one EU country.

China was the third highest country of spam origin in the first half of 2007. Four percent of spam detected by Symantec during this period originated there, compared to six percent in the last half of 2006. This is a continuation of the drop that was first noted in the previous edition of the *Internet Security Threat Report*. This drop may be due to an increasing focus on computer security in China as Internet regulation, such as port blocking by service providers,[183] begins to catch up with its rapid growth and users become more knowledgeable.[184] This can also be seen in the slowing increase of bots in China as discussed in the "Attack Trends" section of this report. It is reasonable to speculate that this could be because some companies that do not do business in China automatically block all email originating there.

Image spam

For the first time, in this volume of the *Internet Security Threat Report*, Symantec is assessing the percentage of spam that is image spam. Image spam is a spam email that does not use text in the body of the message to convey its message; instead, it uses an image embedded in the email. This enables the spam to evade blocking techniques, such as Bayesian filtering,[185] that rely on words in the body of the email. Other methods for detecting spam, such as comparing the MD5 checksum value of known image spam to an incoming email, can also be defeated by making minor changes to the image, such as to the color and size.

During the first half of 2007, 27 percent of all spam blocked by Symantec consisted of image spam (figure 38). While image spam started at a higher level at the beginning of the period, reaching nearly 50 percent of all spam in the first week of January, it showed a marked decline beginning in April and continuing throughout May. The January level is likely due in large part to the rise of the Peacomm Trojan, which sent image spam.[186] While the decline of image spam subsided in June, it did not regain the prominence it achieved at the beginning of the period.

[183] Many Internet service providers block incoming network connections to residential users on certain ports. This can prevent some network worms from propagating, block access to back door servers, and prevent computers from relaying spam.
[184] http://www.networkworld.com/news/2007/012307-china-internet-market-grows-to.html
[185] Bayesian filtering assigns numerical values to certain words that are commonly found in spam. The sum of these values in a particular message is then compared to a score. If the sum exceeds the score, then the message is classified as spam.
[186] http://www.symantec.com/enterprise/security_response/weblog/2007/01/storm_trojan_outbreak_a_spamce.html

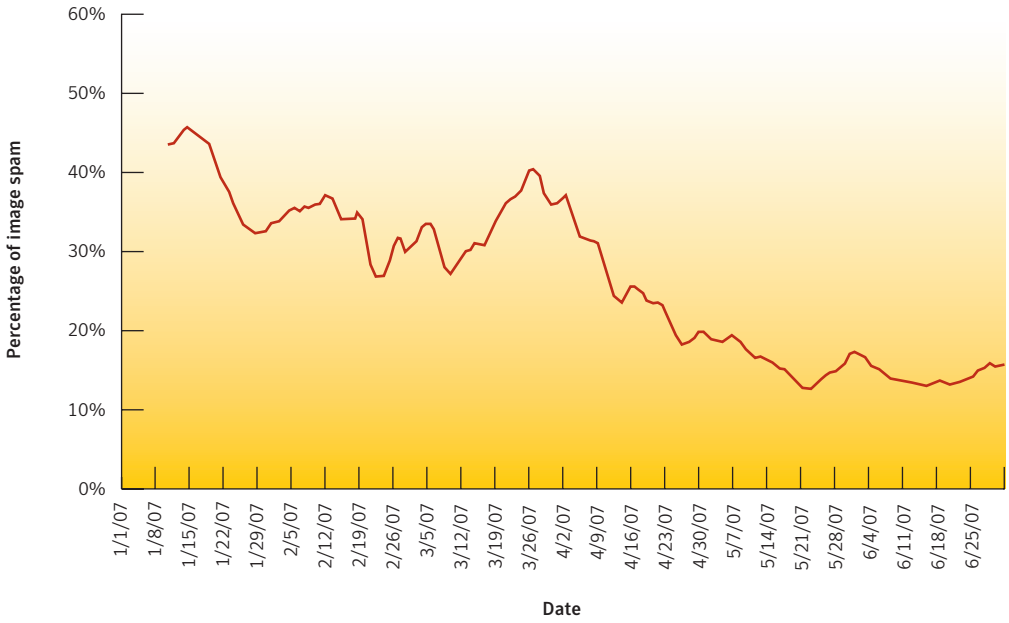

Figure 38. Image spam as a percentage of all spam
Source: Symantec Corporation

The decline in image spam is likely a result of the increased ability of antispam solutions to detect and block it. Also, a vast majority of image spam in the previous reporting was pump and dump stock spam. As was described in the "Top spam categories" of this report, this type of spam has experienced a significant decline, which has likely contributed to the decline of image spam.

Another change of note is that more spam messages are linking to an image hosted on remote servers instead of embedding the image in the message itself. The HTML code in the email will retrieve the image when the user views the message, so the image never passes through antispam filtering. This shows that, as one method of delivering spam loses its effectiveness, spammers will adapt other techniques. Users and network administrators should ensure that their antispam measures are not static—as in the case of many antispam scripts and simple Bayesian filters—and are capable of evolving as attacks change.

Appendix A—Symantec Best Practices

Enterprise Best Practices

1. Employ defense-in-depth strategies, which emphasize multiple, overlapping, and mutually supportive defensive systems to guard against single-point failures in any specific technology or protection method. This should include the deployment of regularly updated antivirus, firewalls, intrusion detection, and intrusion protection systems on client systems.
2. Turn off and remove services that are not needed.
3. If malicious code or some other threat exploits one or more network services, disable or block access to those services until a patch is applied.
4. Always keep patch levels up to date, especially on computers that host public services and are accessible through the firewall, such as HTTP, FTP, mail, and DNS services.
5. Consider implementing network compliance solutions that will help keep infected mobile users out of the network (and disinfect them before rejoining the network).
6. Enforce an effective password policy.
7. Configure mail servers to block or remove email that contains file attachments that are commonly used to spread viruses, such as .VBS, .BAT, .EXE, .PIF, and .SCR files.
8. Isolate infected computers quickly to prevent the risk of further infection within the organization. Perform a forensic analysis and restore the computers using trusted media.
9. Train employees to not open attachments unless they are expected and come from a known and trusted source, and to not execute software that is downloaded from the Internet unless it has been scanned for viruses.
10. Ensure that emergency response procedures are in place. This includes having a backup-and-restore solution in place in order to restore lost or compromised data in the event of successful attack or catastrophic data loss.
11. Educate management on security budgeting needs.
12. Test security to ensure that adequate controls are in place.
13. Be aware that security risks may be automatically installed on computers with the installation of file-sharing programs, free downloads, and freeware and shareware versions of software. Clicking on links and/or attachments in email messages (or IM messages) may also expose computers to unnecessary risks. Ensure that only applications approved by the organization are deployed on desktop computers.

Consumer Best Practices

1. Consumers should use an Internet security solution that combines antivirus, firewall, intrusion detection, and vulnerability management for maximum protection against malicious code and other threats.
2. Consumers should ensure that security patches are up-to-date and that they are applied to all vulnerable applications in a timely manner.
3. Consumers should ensure that passwords are a mix of letters and numbers, and should change them often. Passwords should not consist of words from the dictionary.
4. Consumers should never view, open, or execute any email attachment unless the attachment is expected and the purpose of the attachment is known.
5. Consumers should keep virus definitions updated regularly. By deploying the latest virus definitions, consumers can protect their computers against the latest viruses known to be spreading in the wild.
6. Consumers should routinely check to see if their operating system is vulnerable to threats by using Symantec Security Check at www.symantec.com/securitycheck.
7. Consumers should deploy an antiphishing solution. They should never disclose any confidential personal or financial information unless and until they can confirm that any request for such information is legitimate.
8. Consumers can get involved in fighting cybercrime by tracking and reporting intruders. With Symantec Security Check's tracing service, users can quickly identify the location of potential hackers and forward the information to the attacker's ISP or local police.
9. Consumers should be aware that security risks may be automatically installed on computers with the installation of file-sharing programs, free downloads, and freeware and shareware versions of software. Clicking on links and/or attachments in email messages (or IM messages) may also expose computers to unnecessary risks. Ensure that only applications approved by the organization are deployed on desktop computers.
10. Some security risks can be installed after an end user has accepted the end-user license agreement (EULA), or as a consequence of that acceptance. Consumers should read EULAs carefully and understand all terms before agreeing to them.
11. Consumers should beware of programs that flash ads in the user interface. Many spyware programs track how users respond to these ads, and their presence is a red flag. When users see ads in a program's user interface, they may be looking at a piece of spyware.

Appendix B—Attack Trends Methodology

Attack trends in this report are based on the analysis of data derived from the Symantec Global Intelligence Network, which includes the Symantec DeepSight Threat Management System, Symantec Managed Security Services, the Symantec Honeypot Network, and proprietary Symantec technologies. Symantec combines data derived from these sources for analysis.

Denial of service attacks

Although there are numerous methods for carrying out denial of service (DoS) attacks, Symantec derives this metric by measuring DoS attacks that are carried out by flooding a target with SYN requests.[187] These are often referred to as SYN flood attacks. This type of attack works by overwhelming a target with SYN requests and not completing the initial request, which thus prevents other valid requests from being processed.

In many cases, SYN requests with forged IP addresses are sent to a target, allowing a single attacking computer to initiate multiple connections, resulting in unsolicited traffic, known as backscatter, being sent to other computers on the Internet. This backscatter is used to derive the number of DoS attacks observed throughout the reporting period. Although this methodology will not identify all DoS attacks carried out, it does allow Symantec to assess high-level DoS attack trends.

To determine the countries targeted by DoS attacks, Symantec cross-references the target IP addresses of every detected attack with several third-party, subscription-based databases that link the source IP addresses with the geographic location of the originating computer. While these databases are generally reliable, there is a small margin of error.

Sectors targeted by DoS attacks were identified using the same methodology as targeted countries. However, in this case, the only attackers considered were those carrying out DoS attacks that were detected by IDS and IPS software.

Top originating countries

Symantec identifies the country of origin of attacks by automatically cross-referencing source IP addresses of every attacking IP address with several third-party, subscription-based databases that link the source IP address with the geographic location of the originating computer. While these databases are generally reliable, there is a small margin of error.

Malicious activity by country

To determine the top countries for the “Malicious activity by country” metric, Symantec compiled geographical data on each type of malicious activity to be considered, which included: bot-infected computers, bot command-and-control servers, phishing Web sites, malicious code infections, spam relay hosts, and Internet attacks. The proportion of each activity originating in each country was determined. The mean of the percentages of each malicious activity that originated in each country was calculated. This average determined the proportion of overall malicious activity that originated from the country in question and was used to rank each country.

[187] The TCP protocol requires a three-way exchange to be carried out before any data is sent. The SYN request is the first phase of the three-way exchange. Once a SYN request is received by a server, a SYN-ACK is sent in response. The final step is an ACK response, completing the connection negotiation process.

Symantec also evaluated the top 25 of these countries according to the number of Internet users located there. This measure is meant to remove the bias of high Internet users from the consideration of the "Malicious activity by country" metric. Symantec determined the top 25 countries for network corruption as a percentage of Internet users by using the same data as above. In order to determine this, Symantec divided the amount of malicious activity originating in each of the top 25 countries for malicious activity by the percentage of worldwide Internet users located in that country.

The percentage assigned to each country in the discussion thus corresponds to the proportion of malicious activity that could be attributed to a single (average) Internet user in that country. That is, we first take one average Internet user from each of the top 25 countries and measure their collective malicious activity. The percentage of malicious activity that would be carried out by each person is the proportion assigned to each country.

Malicious activity originating from Fortune 100 organizations

To determine the proportion of malicious activity originating from Fortune 100 organizations, Symantec determined IP address ranges of the Fortune 100 organizations. These IP addresses were determined using autonomous system numbers (ASN). That is, the IP addresses that were registered by the Fortune 100 companies were used to determine the malicious activity originating from them. These IP ranges were in turn used to determine the percentage of malicious activity originating from computers determined to belong to those organizations, including Internet attacks, active bot-infected computers, phishing Web sites, and spam zombies. The attack activity carried out by Fortune 100 companies was compared to the world total to determine the percentage of overall attack activity originating from each organization.

For a number of reasons, the IP addresses used for this analysis may not be exact. For instance, an IP address may be assigned to one company but be used by another. This is particularly true for companies that own many IP addresses. It is also possible for attackers to spoof IP addresses, making it look like their attacks originate from Fortune 100 organizations when they do not. As a consequence, some attacks that are either spoofed or originate from organizations other than Fortune 100 companies may be inadvertently included in this discussion.

Identity theft data breaches

Symantec identifies the proportional distribution of cause and sector for data breaches that may facilitate identity theft based on data provided by Attrition.org.[188] The sector that experienced the loss along with the cause of loss that occurred is determined through analysis of the organization reporting the loss and the method that facilitated the loss.

[188] http://www.attrition.org

Underground economy servers

This metric is based on data that is gathered by proprietary Symantec technologies. These technologies monitor activity on underground economy servers and collect data. Underground economy servers are typically chat servers on which stolen data, such as identities, credit card numbers, access to compromised computers, and email accounts are bought and sold. Each server is monitored by recording communications that take place on them, which typically includes advertisements for stolen data. This data was used to derive the data presented in this metric.

Active bot-infected computers

Symantec identifies bot-infected computers based on coordinated scanning and attack behavior that is observed in network traffic. For an attacking computer to be considered to be participating in coordinated scanning and attacking, it must fit into that pattern to the exclusion of any other activity. This behavioral matching will not catch every bot-infected computer, and may identify other malicious code or individual attackers behaving in a coordinated way as a bot network. This behavioral matching will, however, identify many of the most coordinated and aggressive bot-infected computers. It will also give insight into the population trends of bot-infected computers, including those that are considered to be actively working in a well coordinated and aggressive fashion at some point in time during the reporting period.

Lifespan of bot-infected computers

Using previously identified bot-infected computers, Symantec determined the life span of these infections by measuring the time between their first and last detected activity. However, to ensure that the lifespan reflects a continuous bot infection, if the identified computer was inactive for 30 days or longer it was considered to be disinfected. As such, any further bot-like activity would be considered a new infection.

Bot-infected computers by countries and cities

This metric is based on the same data as “Active bot networks” discussion of the “Attacks Trends” section of the report. Symantec cross-references the IP addresses of every identified bot-infected computer with several third-party subscription-based databases that link the geographic location of systems to IP addresses. While these databases are generally reliable, there is a small margin of error. The data produced is then used to determine the global distribution of bot-infected computers.

Top targeted sectors

For the purposes of the *Internet Security Threat Report*, a targeted attacker is defined as one that is detected attacking at least three users or organizations in a specific sector, to the exclusion of all other sectors. The targeted sector attack rate is a measure of the percentage of all attackers that target only organizations or users in a specific sector and is represented as a proportion of all targeted attacks. Figure 39 represents the proportional sensor distribution for each sector. (Due to rounding of numbers, the cumulative percentage of sensors may exceed 100.) Sectors with less than 10 sensors have been excluded from the resulting totals.

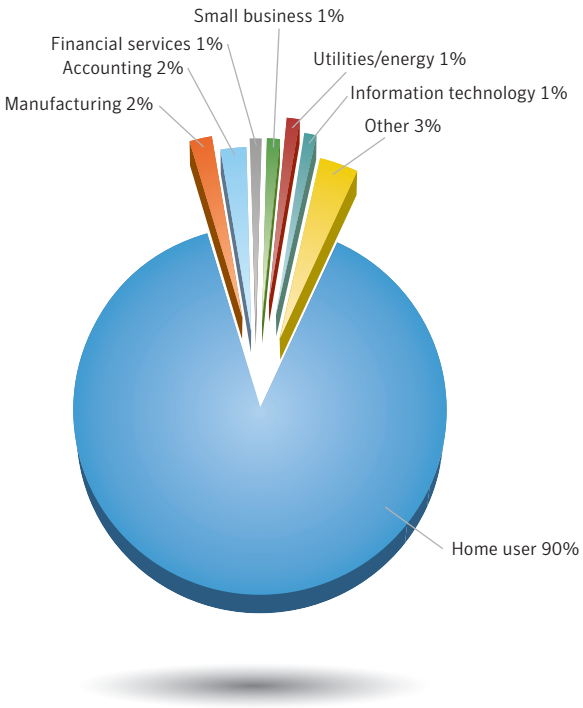

Figure 39. Distribution of sensors by sector
Source: Symantec Corporation

Appendix C—Vulnerability Trends Methodology

The "Vulnerability Trends" section of the Symantec *Internet Security Threat Report* discusses developments in the discovery and exploitation of vulnerabilities over the six-month reporting period and compares it to activity observed in the previous six-month period. This section will discuss the methods by which the data was gathered and analyzed to come to the conclusions that are presented in the "Vulnerability Trends" section.

Symantec maintains one of the world's most comprehensive databases of security vulnerabilities, consisting of over 22,000 distinct entries. Each distinct entry is created and maintained by Symantec threat analysts who vet the content for accuracy, veracity, and the applicability of its inclusion in the vulnerability database based on available information. The following metrics are based on the analysis of that data by Symantec researchers:

- Total number of vulnerabilities disclosed
- Severity of vulnerabilities
- Web application vulnerabilities
- Easily exploitable vulnerabilities
- Operating system patch development time
- Web browser vulnerabilities
- Zero-day vulnerabilities
- Database vulnerabilities
- Unpatched enterprise vulnerabilities

The ways in which the data for the remaining metrics is gathered and analyzed will be discussed in the remainder of this methodology.

Vulnerability classifications

Following the discovery and/or disclosure of a new vulnerability, Symantec analysts gather all relevant characteristics of the new vulnerability and create an alert. This alert describes important traits of the vulnerability, such as the severity, ease of exploitation, and a list of affected products. These traits are subsequently used both directly and indirectly for this analysis.

Vulnerability type

After discovering a new vulnerability, Symantec threat analysts classify the vulnerability into one of 12 possible categories based on the available information. These categories focus on defining the core cause of the vulnerability, as opposed to classifying the vulnerability merely by its effect.

The classification system is derived from the academic taxonomy presented by Taimur Aslam, et al (1996) to define classifications of vulnerabilities.[189]

[189] "Use of a Taxonomy of Security Faults" http://ftp.cerias.purdue.edu/pub/papers/taimur-aslam/aslam-krsul-spaf-taxonomy.pdf

Possible values are indicated below; the previously mentioned white paper provides a full description of the meaning behind each classification:

- Boundary condition error
- Access validation error
- Origin validation error
- Input validation error
- Failure to handle exceptional conditions
- Race condition error
- Serialization error
- Atomicity error
- Environment error
- Configuration error
- Design error

Severity of vulnerabilities

Severity of vulnerabilities has been discussed in previous versions of the Symantec *Internet Security Threat Report*, however, it was omitted in Volume X of the report (September 2006) because Symantec's adoption of the Common Vulnerability Scoring System (CVSS) V1.0.[190]

The "Severity of vulnerabilities" metric that has been included in this report corresponds to the base score field of the CVSS. The base score is representative of the inherent properties of a vulnerability, such as:

- The degree of confidentiality, integrity, or availability of data that may be affected by the vulnerability
- Local versus remote exploitability
- Whether or not authentication is required for exploitation
- If there are additional factors that may complicate exploitation of the vulnerability

These values are not adjusted for temporal factors such as the availability of exploit code. The base score is intended to be a static value that should only change if additional information is made available that changes the inherent characteristics of the vulnerability. The base score can have a value of zero to 10.

For the sake of categorizing vulnerabilities by their respective severities, the following standard is used:

- **Low severity (base score of 0–3):** Successful exploitation of these vulnerabilities will have a minimal impact on the confidentiality, integrity, and availability of data stored upon or transmitted over systems on which the vulnerability may be found. These vulnerabilities also tend to be local in nature, have a high degree of access complexity, and may require authentication to be exploited successfully.
- **Medium severity (base score of 4–7):** Successful exploitation of these vulnerabilities could allow a partial compromise of the confidentiality, integrity, and availability of data stored upon or transmitted over systems on which the vulnerability may be found, although this may not always be the case. These vulnerabilities can be exploited remotely over a network and may have a lower access complexity or may or may not require authentication to successfully exploit.

[190] http://www.first.org/cvss/v1

- **High severity (base score of 8–10):** These vulnerabilities have innate characteristics that present the highest threat profile. Successful exploitation often allows a complete compromise of the confidentiality, integrity, and availability of data stored upon or transmitted over systems on which the vulnerability may be found. These vulnerabilities are exploited remotely across a network, have a low degree of access complexity, and usually do not require authentication prior to successful exploitation.

Base scores are computed from related fields in the Symantec Vulnerability Database. They are then categorized into low, medium, and high, as described above, and broken out by reporting period.

Easily exploitable vulnerabilities

The "Easily exploitable vulnerabilities" metric covers vulnerabilities that attackers can exploit with little effort based on publicly available information. The vulnerability analyst assigns an exploit availability rating after thoroughly researching the need for and availability of exploits for the vulnerability.

The "Easily exploitable vulnerabilities" metric replaces the "Ease of exploitation" metric, which was included in the *Internet Security Threat Report* prior to Volume XI (March 2007). This change was made to accommodate Symantec's adoption of the exploitability rating in the CVSS.

All vulnerabilities are classified into one of four possible categories defined by the CVSS, as described below:

1. **Unconfirmed:** Would-be attackers must use exploit code to make use of the vulnerability; however, no such exploit code is publicly available.
2. **Proof-of-concept:** Would-be attacks must use exploit code to make use of the vulnerability; however, there is only proof-of-concept exploit available that is not functional enough to fully exploit the vulnerability.
3. **Functional:** This rating is used under the following circumstances:
   - Exploit code to enable the exploitation of the vulnerability is publicly available to all would-be attackers; and/or,
   - Would-be attackers can exploit the vulnerability without having to use any form of exploit code;
   - In other words, the attacker does not need to create or use complex scripts or tools to exploit the vulnerability.
4. **High:** The vulnerability is reliably exploitable and there have been instances of self-propagating malicious code exploiting the vulnerability in the wild.

For the purposes of this report, the last two categories of vulnerabilities are considered "easily exploitable" because the attacker requires only limited sophistication to exploit the vulnerability. The first two categories of vulnerability are considered more difficult to exploit because attackers must develop their own exploit code or improve an existing proof-of-concept to make use of the vulnerability.

Window of exposure for enterprise vendors

Symantec records the time lapse between the publication of an initial vulnerability report and the appearance of third-party exploit code; this is known as the exploit development time. The time period between the disclosure date of a vulnerability and the release date of an associated patch is known as the patch development time.[191] The time lapse between the public release of exploit code and the time that the affected vendor releases a patch for the affected vulnerability is known as the window of exposure.

The average window of exposure is calculated as the difference in days between the average exploit development time and the average patch development time. (Explanations of the exploit development time average and the patch development time average are included below.) During this time, the computer or system on which the affected application is deployed may be susceptible to attack, as administrators have no official recourse against the vulnerability and must resort to best practices and workarounds to reduce the risk of exploitation.

It is important to note that the set of vulnerabilities included in this metric is limited and does not represent all software from all possible vendors. Instead, it only includes vendors who are classified as enterprise vendors. The purpose is to illustrate the window of exposure for widely deployed mission-critical software. Because of the large number of vendors with technologies that have a very low deployment (which form the majority), only exploits for technologies from enterprise vendors (that is, those that generally have widespread deployment) are included. Vulnerabilities in those vendors' products will likely affect more enterprises than those in less widely deployed technologies. Those vendors are:

- CA (Computer Associates)
- Cisco
- EMC
- HP
- IBM
- McAfee
- Microsoft
- Oracle
- Sun
- Symantec

Patch development time for enterprise vendors

The patch development time is the time period between the disclosure date of a vulnerability and the release date of an associated patch. Only those patches that are independent objects (such as fixes, upgrades, etc.) are included in this analysis. Other remediation solutions—such as workaround steps, for instance—are excluded.

For each individual patch from these vendors, the time lapse between the patch release date and the publish date of the vulnerability is computed. The mean average is calculated from the aggregate of these. As some vendors may release more patches than others for a particular vulnerability, Symantec considers only the first instance of a single patch for each vulnerability. This metric is incorporated when computing the window of exposure, which is calculated as the difference between the average patch development time and the average exploit development time.

[191] This statistic only considers specific file-based patches or upgrades, and not general solutions. Instances in which the vendor provides a workaround or manual fix, for example, are not included.

Exploit code development time for enterprise vendors

The ability to measure exploit code development time is limited and applies only to vulnerabilities that would normally require exploit code. Therefore, the metric is based on vulnerabilities that Symantec considers to be of sufficient complexity, and for which functional exploit code was not available until it was created by a third party. This consideration, therefore, excludes the following:

- Vulnerabilities that do not require exploit code (unconfirmed exploitability);
- Vulnerabilities associated with non-functional proof-of-concept code (proof-of-concept exploitability).

The date of vulnerability disclosure is based on the date of the first publicly available reference (such as a mailing-list post). The date of exploit code publication is the date of the first publicly known reference to the exploit code. Because the purpose of this metric is to estimate the time it takes for exploit code to materialize as a result of active development, exploit code publication dates that fall outside of the 30-day range from initial vulnerability publication are excluded from this metric. It is assumed that exploit code that was published after this period was not actively developed from the initial announcement of the vulnerability.

Because this metric only considers the appearance of the first functional exploit, it is possible that reliable exploits that improve upon the initial exploit may appear later. These exploits may take much longer to develop, but are not considered because the window of exposure begins as soon as the first functional exploit surfaces.

The time lapse between the disclosure of a vulnerability and the appearance of exploit code for that vulnerability is determined. The aggregate time for all vulnerabilities is determined and the average time is calculated. This metric is incorporated when computing the window of exposure, which is the difference between the average patch development time and the average exploit development time.

Operating system patch development time

This metric has a similar methodology to the “Patch development time for enterprise vendors” metric, which was explained earlier in this methodology. However, instead of applying it to enterprise-scale vendors, the patch development time average is calculated from patched vulnerabilities for the following operating systems:

- Apple Mac OS X
- Hewlett-Packard HP-UX
- Microsoft Windows
- Red Hat Linux (including enterprise versions and Red Hat Fedora)
- Sun Microsystems Solaris

The sample set includes only vulnerabilities that are considered medium severity or higher, based on their CVSS base score. An average is calculated from the patch release times for each vulnerability in the reporting period per operating system. The patch development time average for each operating system is then compared.

Operating system time to patch by type

This is an analysis of the patched vulnerabilities in the data set for the "Operating system patch development time" metric. For each vendor studied in that metric, each vulnerability is divided into one of the following categories:

- **Browser vulnerabilities:** These vulnerabilities threaten Web browser applications through remote attack vectors.
- **Client-side vulnerabilities:** These vulnerabilities threaten network client applications or non-networked applications that process malicious data that may arrive through another networked application. Remote attack vectors may exist, but client-side vulnerabilities usually require some amount of user interaction on the part of the victim to be exploited.
- **Local vulnerabilities:** These are vulnerabilities that require local access in order to be successfully exploited. Local attacks may affect a large variety of applications that may or may not include network capabilities. The differentiator is that these vulnerabilities are not exploitable by remote attackers unless they can log on to the system and run commands as an unprivileged user.
- **Server vulnerabilities:** These are vulnerabilities that affect server applications. Server applications are typically defined as applications that are accessible to remote clients via connections on a range of TCP/UDP ports. Server vulnerabilities generally do not require user interaction on the part of the victim beyond enabling and starting the service so that it listens for incoming requests.
- **Other:** These are vulnerabilities that do not fall discretely into any of the previous categories. They can include applications for which the distinction is blurred between server and client, or hardware platforms in which the affected component cannot be described by any of the other categories.

These categories are generally defined by the attack vector and by the type of application that is affected. The specific categories were devised so that the majority of vulnerabilities could easily be classified within them, with little overlap between categories, so that the total percentage of all categories would equal 100 percent.

Window of exposure for Web browsers

This metric has a similar methodology to the "Window of exposure for enterprise vendors" metric. However, instead of applying it to enterprise-scale vendors, the window of exposure is calculated for vulnerabilities associated with the following Web browsers:

- Apple Safari
- Microsoft Internet Explorer
- Mozilla Firefox and Mozilla browsers
- Opera

Symantec records the window of time between the publication of an initial vulnerability report and the appearance of third-party exploit code; this is known as the exploit code development time. The time period between the disclosure date of a vulnerability and the release date of an associated patch is known as the patch development time. The time lapse between the public release of exploit code and the time that the affected vendor releases a patch for the affected vulnerability is known as the window of exposure.

The average window of exposure is calculated as the difference in days between the average patch development time and the average exploit code development time. During this time, the computer or system on which the affected application is deployed may be susceptible to attack, as administrators may have no official recourse against a vulnerability and must resort to best practices and workarounds to reduce the risk of attacks. Explanations of the average exploit development time and the average patch development time are included below.

Patch development time for Web browsers

The cumulative patch development time for vulnerabilities affecting each browser is calculated. Each cumulative time is then divided by the number of vulnerabilities affecting that browser to determine the average patch development time for that browser. The patch development time average for each browser is then compared. This metric is used to compute the window of exposure for Web browsers, which amounts to the difference between the average patch development time and the average exploit code development time.

Exploit code development time for Web browsers

The cumulative exploit code development time for each vulnerability affecting a Web browser is calculated. Each cumulative time is then divided by the number of vulnerabilities affecting that browser to determine the average exploit code development time for that browser. The exploit development time average for each browser is then compared. This metric is used to compute the window of exposure, which amounts to the difference between the average patch development time and the average exploit code development time.

Web browser vulnerabilities

This metric will offer a comparison of vulnerability data for numerous Web browsers, namely: Microsoft Internet Explorer, the Mozilla browsers (which includes Firefox), Opera, and Safari. However, in assessing the comparative data, the following important caveats should be kept in mind before making any conclusions:

- The total number of vulnerabilities in the aforementioned Web browsers was computed for this report;
- This includes vulnerabilities that have been confirmed by the vendor and those that are not vendor confirmed.

Previous versions of the *Internet Security Threat Report* have discussed vulnerabilities according to whether they were vendor confirmed or non-vendor confirmed because vulnerabilities that were not confirmed were also included in the data. This differentiation was important, especially given the disparity in patch times between vendors. However, starting with Volume X of the *Internet Security Threat Report*, this convention was no longer followed. This version of the report does not differentiate between vendor-confirmed vulnerabilities and non-vendor-confirmed vulnerabilities when calculating the total number of vulnerabilities.

Individual browser vulnerabilities are notoriously difficult to pinpoint and identify precisely. A reported attack may be a combination of several conditions, each of which could be considered a vulnerability in its own right. This may distort the total vulnerability count. Some browser issues have also been improperly identified as operating system vulnerabilities or vice versa. This is, in part, due to increasing operating system integration that makes it difficult to correctly identify the affected component in many cases.

- Many vulnerabilities in shared operating system components can be exposed to attacks through the browser. This report enumerates only those vulnerabilities that are known to affect the browser itself where sufficient information is available to make the distinction.
- Not every vulnerability that is discovered is exploited. As of this writing, there has been no widespread exploitation of any browser except Microsoft Internet Explorer. This is expected to change as other browsers become more widely deployed.

Browser plug-in vulnerabilities

Browser plug-ins are technologies that extend the functionality of the Web browser. They may be developed by the vendor or by a third party. Some plug-ins provide support for additional application programming languages or environments, such as Java or Flash. Others are applications in their own right that run in the browser. Examples of these include ActiveX objects for Internet Explorer, Firefox extensions, or Opera widgets.

This metric enumerates publicly documented vulnerabilities that affect browser plug-ins. These vulnerabilities are further classified, when applicable, into general groups of browser plug-in technologies.

Symantec makes an effort to identify all vulnerabilities affecting the various classes of browser plug-in. Vulnerabilities that affect the browser itself are not included in the data for this metric when it is possible to make this distinction. In cases where a Web browser ships with a particular plug-in, vulnerabilities affecting that plug-in will be counted. Although in this case, the plug-in may be included in the default browser installation, it is still considered a separate technology and not a native feature of the browser.

Native feature are considered to be features intrinsic to the primary function of the browser such as support for HTTP/HTTPS, HTML rendering, JavaScript, and other standards that are commonly implemented in most Web browsers. Technologies such as Java and Flash may be common to many Web browsers but they are intended to extend their functionality to support additional types of content and are typically optional components.

The definition of browser plug-in for this report is limited to technologies that are hosted on the same computer as the browser, and whose installation and configuration is managed through the browser or operating system. This distinguishes them from content that is intended to run inside the browser but is typically external to the browser such as Java applets or Flash movies. This content is rendered or executed by a browser plug-in but is not considered to be a plug-in in its own right.

Zero-day vulnerabilities

This metric quantifies the number of zero-day vulnerabilities that have been documented during the relevant reporting periods of the current *Internet Security Threat Report*. For the purpose of this metric, a zero-day vulnerability is one for which there is sufficient public evidence to indicate that the vulnerability has been exploited in the wild prior to being publicly known. It may not have been known to the vendor prior to exploitation, and the vendor had not released a patch at the time of the exploit activity.

This metric is derived from public sources and the Symantec vulnerability database. This metric is meant to calculate the number of high-profile, publicly documented zero-day vulnerability instances during the relevant reporting periods.

Database vulnerabilities

This metric offers a comparison of the vulnerabilities across multiple database vendors and implementations. For the purpose of this report, databases to be assessed were chosen to reflect the most widely deployed database implementations and to compare commercial and open source vendors.[192] To this end, the following five database implementations are discussed:

- IBM® DB2®
- Microsoft SQL Server
- MySQL
- Oracle
- PostgreSQL

The volume of database vulnerabilities is determined by querying the vulnerability database for vulnerabilities that affect the aforementioned database implementations. The results are broken out by implementation and reporting period.

Unpatched enterprise vulnerabilities

Unpatched vulnerabilities are vulnerabilities that have no vendor remediation at the time that data for the report was collected.[193] This metric tracks the number of unpatched vulnerabilities affecting enterprise-scale technologies. Individual vendors are identified and correlated with the number of unpatched vulnerabilities affecting them. It is possible that some vendors will have no vulnerabilities affecting them during a given reporting period or that none of the vulnerabilities affecting them are considered unpatched.

[192] Oracle, DB2, and Microsoft SQL Server are the three most widely deployed commercial database implementations (http://databases.about.com/b/a/016881.htm). MySQL and PostgreSQL are the two most popular open-source databases (http://www.mysql.com/why-mysql/marketshare).
[193] For the purpose of this report, patched vulnerabilities are those with vendor-supplied patches or upgrades. Vendor-supplied or third-party workarounds are not counted as patches.

The status of some vulnerabilities may have changed since data was collected; vendors may have released patches for vulnerabilities included in the data set and new vulnerabilities may have been published that are considered unpatched. The nature of unpatched vulnerabilities means that the data may include vulnerabilities that are unverified and may have been reported by a single source with no other corroboration. However, the data also includes vulnerabilities that have been acknowledged but not fixed by the vendor. In rare instances, the legitimacy of a vulnerability may be in dispute, but in all such cases these disputes remain unresolved at the time of data collection. Symantec excludes all vulnerabilities that are provably false from this and other metrics in the report.

It is also important to note that the set of vulnerabilities included in this metric is limited and does not represent all software from all possible vendors. Instead, it only includes vendors who are classified as enterprise vendors. The purpose is to illustrate the window of exposure for widely deployed mission-critical software. Because of the large number of vendors with technologies that have a very low deployment (which form the majority), only exploits for technologies from enterprise vendors (that is, those that generally have widespread deployment) are included. Vulnerabilities in those vendors' products will likely affect more enterprises than those in less widely deployed technologies. Those vendors are:

- CA (Computer Associates)
- Cisco
- EMC
- HP
- IBM
- McAfee
- Microsoft
- Oracle
- Sun
- Symantec

Appendix D—Malicious Code Trends Methodology

The trends in the "Malicious Code Trends" section are based on statistics from malicious code samples reported to Symantec for analysis. Symantec gathers data from over 120 million client, server, and gateway systems that have deployed Symantec's antivirus products in both consumer and corporate environments. The Symantec Digital Immune System and Scan and Deliver technologies allow customers to automate this reporting process. Observations in the "Malicious Code Trends" section are based on empirical data and expert analysis of this data. The data and analysis draw primarily from two databases described below.

Infection database

To help detect and eradicate computer viruses, Symantec developed the Symantec AntiVirus Research Automation (SARA) technology. Symantec uses this technology to analyze, replicate, and define a large subset of the most common computer viruses that are quarantined by Symantec Antivirus customers.

On average, SARA receives hundreds of thousands of suspect files daily from both enterprise and individual consumers located throughout the world. Symantec then analyzes these suspect files, matching them with virus definitions. An analysis of this aggregate data set provides statistics on infection rates for different types of malicious code.

Malicious code database

In addition to infection data, Symantec Security Response analyzes and documents attributes for each new form of malicious code that emerges both in the wild and in a "zoo" (or controlled laboratory) environment. Descriptive records of new forms of malicious code are then entered into a database for future reference. For this report, a historical trend analysis was performed on this database to identify, assess, and discuss any possible trends, such as the use of different infection vectors and the frequency of various types of payloads.

In some cases, Symantec antivirus products may initially detect new malicious code heuristically or by generic signatures. These may later be reclassified and given unique detections. Because of this, there may be slight variance in the presentation of the same data set from one volume of the *Internet Security Threat Report* to the next.

Geographic location of malicious code instances

Several third-party subscription-based databases that link the geographic locations of systems to IP addresses are used along with proprietary Symantec technology to determine the location of computers reporting malicious code instances. While these databases are generally reliable, there is a small margin of error. The data produced is then used to determine the global distribution of malicious code instances.

Previously unseen malicious code threats

This metric derives its data from the Symantec Honeypot Network. Computers compromised on the honeypot network track and analyze each piece of malicious code that is installed by the attacker. Symantec defines previously unseen malicious threats as those that have not been installed by attackers on the Symantec Honeypot Network. The proportion of previously unseen malicious code threats is derived by comparison with the total number of distinct malicious code threats observed.

Percentage of malicious code that exploits vulnerabilities

Symantec maintains a malicious code database to analyze and document individual instances of malicious code. This database contains 8,000 distinct entries, with the earliest discovery dating back to 1998. The database includes metadata for classifying malicious code by type, discovery date, and by threat profile, in addition to providing mitigating factors and manual removal steps. Where applicable, this database includes correlations between malicious code instances and vulnerabilities from the Symantec vulnerability database. This capability was used as a basis for the data in this metric. Symantec examined the means by which the malicious code propagated, and counted those that propagate by exploiting vulnerabilities.

Appendix E—Phishing and Spam Trends Methodology

Traditionally, the Symantec *Internet Security Threat Report* has broken security threats down into three general categories: attacks, vulnerabilities, and malicious code. However, as Internet-based services and applications have expanded and diversified, the potential for computer programs to introduce other types of security risks has increased. The emergence of new risks, particularly spam and phishing has necessitated an expansion of the traditional security taxonomy.

Symantec has monitored these new concerns as they have developed. In particular, the *Internet Security Threat Report* assesses these risks according to two categories: phishing and spam. The methodology for each of these discussions will be discussed in the sections below.

Phishing

Phishing attack trends in this report are based on the analysis of data derived from the Symantec Probe Network. Symantec Brightmail AntiSpam data is assessed to gauge the growth in phishing attempts as well as the percentage of Internet mail that is determined to be phishing attempts. Symantec Brightmail AntiSpam field data consists of statistics reported back from customer installations that provide feedback about the detection behaviors of antifraud filters as well as the overall volume of mail being processed.

It should be noted that different monitoring organizations use different methods to track phishing attempts. Some groups may identify and count unique phishing messages based solely on specific content items such as subject headers or URLs. These varied methods can often lead to differences in the number of phishing attempts reported by different organizations.

Phishing attempt definition

The Symantec Probe Network is a system of over two million decoy accounts that attract email messages from 20 different countries around the world. It encompasses more than 600 participating enterprises and attracts email samples that are representative of traffic that would be received by over 250 million mailboxes. The Probe Network covers countries in the Americas, Europe, Asia, Africa, and Australia/Oceania.

The Symantec Probe Network data is used to track the growth in new phishing activity. A phishing attempt is a group of email messages with similar properties, such as headers and content, that is sent to unique users. The messages attempt to gain confidential and personal information from online users.

Symantec Brightmail AntiSpam software reports statistics to Symantec Security Response that indicate messages processed, messages filtered, and filter-specific data. Symantec has classified different filters so that spam statistics and phishing statistics can be determined separately. Symantec Brightmail AntiSpam field data is used to identify general trends in phishing email messages.

Explanation of research inquiries

This section will provide more detail on specific methodologies used to produce the data and statistics in this report. While most methodologies are adequately explained in the analysis section of the report, the following investigations warrant additional detail.

Unique phishing messages

Symantec maintains automated systems to identify new unique phishing messages received by the Symantec Probe Network. Messages are grouped into attacks based on similarities in the message bodies and headers. Sample messages are then passed through general fraud heuristics to identify messages as potential phishing attempts. Symantec reviews events that are identified as phishing attempts for the purposes of confirmation and to develop filters for those messages.

The data presented in this section is based on monthly totals in the number of new unique phishing messages discovered and ruled upon by Symantec Security Response. Security Response addresses only those phishing messages not caught by existing antispam and antifraud filters. Existing filters refer only to those antispam and antifraud filters used across the Symantec Brightmail AntiSpam customer base.

Some phishing messages will be captured in the field based upon predictive filters (heuristics); however, not all of Symantec's customers utilize this technology or have upgraded to this technology. Therefore, the messages are still reviewed by Security Response for development of filters that are more widely dispersed.

Blocked phishing attempts

The number of blocked phishing attempts is calculated from the total number of phishing email messages that were blocked in the field by Symantec Brightmail AntiSpam antifraud filters. The data for this section is based on monthly totals.

Phishing activity by sector

The Symantec Phish Report Network is an extensive antifraud community in which members contribute and receive fraudulent Web site addresses for alerting and filtering across a broad range of solutions. These sites are categorized according to the brand being phished and the industry to which it belongs. The Phish Report Network has members and contributors that send in phishing attacks from many different sources. This includes a client detection network that detects phishing Web sites as the clients visit various Web sites on the Internet. It also includes server detection from spam emails.

The sender confirms all spoofed Web sites before sending the address of the Web site into the Phish Report Network. After the spoofed site is sent into the Phish Report Network, Symantec spoof detection technology is used to verify that the Web site is a spoof site. Research analysts manage the Phish Report Network console 24 hours a day, 365 days of the year, and manually review all spoof sites sent into the Phish Report Network to eliminate false positives.

Top countries hosting phishing Web sites

The data for this section is determined by gathering links in phishing email messages and cross-referencing the addresses with several third-party subscription-based databases that link the geographic locations of systems to IP addresses. In this case, Symantec counts phishing Web sites as the number of unique IP addresses hosting Web pages used for phishing. While these databases are generally reliable, there is a small margin of error. The data produced is then used to determine the global distribution of phishing Web sites.

Phishing Web sites hosted on the same IP address

This data is compiled by comparing phishing URLs to the IP addresses to which they resolve. The number of URLs resolving to the same addresses are then calculated to determine the number of sites hosted.

Automated phishing toolkits

The data in this section is derived from URLs gathered by the Symantec Phish Report Network. The URLs are sorted and grouped according to specific patterns indicating they were generated by an automated script or phishing kit. Each phishing kit generates URLs with a distinct signature and can be grouped according to these distinguishing characteristics. The monthly total of each group of URLs indicates the level of use of each automated phishing kit.

Core brands being phished

For each phishing Web site Symantec observed during this period, the date and time of the detection was noted along with the name of the brand being spoofed by the Web site. The brand being spoofed is identified using a combination of automated tools and assessment by a Symantec analyst.

Core brands were determined by identifying six lists of brands, one for each month from January through June 2007, in which a new Web site spoofing that brand was reported. The core brands, then, are those that were present on each of these lists. In other words, the core brands were those for which a new phishing Web site was known to have been created in each month of this reporting period.

Spam

The Symantec Probe Network is a system of over two million decoy accounts that attract email messages from 20 different countries around the world. It encompasses more than 600 participating enterprises and attracts email samples that are representative of traffic that would be received by over 250 million mailboxes. The Probe Network includes accounts in countries in the Americas, Europe, Asia, Africa, and Australia/Oceania.

Spam trends in this report are based on the analysis of data derived from both the Symantec Probe Network as well as Symantec Brightmail AntiSpam field data. Symantec Brightmail AntiSpam software reports statistics to the Brightmail Logistical Operations Center (BLOC) indicating messages processed, messages filtered, and filter-specific data.

Symantec has classified different filters so that spam statistics and phishing statistics can be determined separately. Symantec Brightmail AntiSpam field data includes data reported back from customer installations providing feedback from antispam filters as well as overall mail volume being processed.

Symantec Brightmail AntiSpam only gathers data at the SMTP layer and not the network layer, where DNS block lists typically operate. This is because SMTP-layer spam filtering is more accurate than network-layer filtering and is able to block spam missed at the network layer. Network layer-filtering takes place before email reaches the enterprise mail server. As a result, data from the SMTP layer is a more accurate reflection of the impact of spam on the mail server itself.

Sample set normalization

Due to the numerous variables influencing a company's spam activity, Symantec focuses on identifying spam activity and growth projections with Symantec Brightmail AntiSpam field data from enterprise customer installations having more than 1,000 total messages per day. This normalization yields a more accurate summary of Internet spam trends by ruling out problematic and laboratory test servers that produce smaller sample sets.

Explanation of research inquiries

This section will provide more detail on specific methodologies used to produce the data and statistics in this report. While most methodologies are adequately explained in the analysis section of the report, the following investigations warranted additional detail.

Spam as a percentage of email scanned

The data for this section is determined by dividing the number of email messages that trigger antispam filters in the field by the total number of email messages scanned. These filters are distributed across the Symantec Brightmail AntiSpam customer base. The data for this section is based on monthly totals.

Top ten countries of spam origin

The data for this section is determined by calculating the frequency of originating server IP addresses in email messages that trigger antispam filters in the field. The IP addresses are mapped to their host country of origin and the data is summarized by country based on monthly totals. The percentage of spam per country is calculated from the total spam detected in the field.

It should be noted that the location of the computer from which spam is detected being sent is not necessarily the location of the spammer. Spammers can build networks of compromised computers globally and thereby use computers that are geographically separate from their location.

Top countries by spam zombies

The data in this section is determined by examining the IP addresses in spam messages received by the Symantec Probe Network. IP addresses that meet a certain volume requirement are processed through a set of heuristics to determine if they are behaving like zombie servers. If an IP address meets some or all of the heuristic requirements, it will be listed as a zombie IP address. Symantec then cross-references the addresses with several third-party subscription-based databases that link the geographic locations of systems to IP addresses. While these databases are generally reliable, there is a small margin of error. The data produced is then used to determine the global distribution of spam zombies.

About Symantec

Symantec is a global leader in infrastructure software, enabling businesses and consumers to have confidence in a connected world. The company helps customers protect their infrastructure, information, and interactions by delivering software and services that address risks to security, availability, compliance, and performance. Headquartered in Cupertino, Calif., Symantec has operations in 40 countries. More information is available at www.symantec.com.

For specific country offices and contact numbers, please visit our Web site. For product information in the U.S., call toll-free 1 (800) 745 6054.

Symantec Corporation
World Headquarters
20330 Stevens Creek Blvd.
Cupertino, CA 95014 USA
+1 (408) 517 8000
1 (800) 721 3934
www.symantec.com

09/07 12755151