IBM Global Technology Services
August 2009

IBM Internet Security Systems™ X-Force® 2009 Mid-Year Trend and Risk Report

Table of Contents

Overview

The IBM Internet Security Systems X-Force® research and development team discovers, analyzes, monitors and records a wide array of computer security threats and vulnerabilities. According to X-Force observations, many new and surprising trends surfaced throughout the first half of 2009. We hope that the information presented in this report about these trends will provide a foundation for planning your information security efforts for the rest of 2009 and beyond.

2009 Mid-Year Highlights

Vulnerabilities

- *The number of new vulnerability disclosures in the first half of the year is at the lowest level in the past four years, and the number of new, high severity vulnerability disclosures is down by nearly 30 percent in comparison to last year. These changes are mainly driven by declines in two major categories of Web application vulnerabilities (SQL injection and file include) and the top category of client-side vulnerabilities (ActiveX® controls).*
- *Microsoft®, after three years of holding the top spot of vendor with the most vulnerability disclosures, has dropped down to number three. Sun®, who broke the top five for the first time in 2008, has taken Microsoft's place as the vendor with the most vulnerability disclosures so far this year. However, looks can be deceiving—Sun's shift is most likely due to a change in vulnerability disclosure policy, not a change in overall software quality.*
- *Drupal and Joomla! are the only Web application vendors that remain from 2008 in the top ten vendors with the most disclosures list. When it comes to unpatched vulnerabilities, Joomla! tops the charts, and Drupal and TYPO3 also show up in the list.*
- *As for operating systems, Apple® was narrowly surpassed by Sun Solaris in the first half of this year for new operating system disclosures. However, Microsoft is number one if you only consider the critical and high disclosures.*

Exploitation

- *For Web applications, cross-site scripting attacks appear to be declining while injection attacks continued to see significant increases. After leveling off briefly in December of 2008, SQL injection attacks spiked again this Spring, jumping 46 percent in April and then another 76 percent in May.*
- *For clients, vulnerability disclosures and exploits targeted at document readers like Office applications and PDF files skyrocketed in the first half of this year as did their obfuscation, making it harder and harder to block in-the wild exploits. A PDF vulnerability made the top five exploitation list for the first time.*
- *New Firefox vulnerability disclosures surpassed the number of new Internet Explorer disclosures. ActiveX disclosures are continuing to slow, but are still the predominately exploited type of client-side vulnerability and the largest category of new vulnerability disclosures affecting clients.*

Malware and the Malicious Web

- *The number of new malicious Web links discovered in the first half of 2009 increased by 508 percent in comparison to the first half of 2008. Although the majority of these links are hosted on malicious servers located in China and the United States, the overall number of countries with at least one malicious link has significantly increased, up 80 percent over the entire year of 2008.*
- *Malicious Web sites continue to flourish, but so are other techniques of enticing users to click on malicious links. In addition to spam and phishing, cyber criminals are finding ways to have legitimate (or seemingly legitimate) Web sites host links to their malware.*
- *In addition to some "seemingly legitimate" categories you might expect like Gambling and Pornography, Search Engines and Social Media Web sites like blogs and bulletin boards are also in the top categories of Web sites compromised or simply abused by attackers to host malicious links.*
- *Trojans continue to take up an even greater percentage of the new malware discovered this year. They have gained nine percentage points, comprising 55 percent of all the new malware discovered in the first half of this year in comparison to 46 percent in 2008.*
- *The monetary investment and sophistication in malware in the first half of 2009 was unprecedented, and we have much to learn if we are to stop as bad as or worse than the kind of threat we saw in Conficker.*

Spam and Phishing

- *Although URL spam is still the predominate type of spam, its usage slowed slightly in the first half of this year, and image-based spam made a comeback after practical extinction in 2008.*
- *Spam levels took longer to recover after McColo than originally thought. Finally, in May of this year, the levels finally reached (and surpassed) the level seen just before the shutdown.*
- *Spammers are increasingly using trusted domain URLs in spam messages, and in the first half of this year, the number of trusted domains (corporate or other "official" Web sites not directly controlled by spammers)seen in spam messages surpassed the number of domains set up by spammers specifically for spam. Trusted domains are often used as a decoy in spam (to fool end-users and spam filters) and are sometimes abused by spammers when they put their spam messages in areas of trusted Web sites that allow anonymous postings.*
- *Phishing decreased dramatically in the first half of this year due to the shift away from financial targets. Analysts believe that banking Trojans are taking the place of financial targets that were typically phished in the past. Online payment targets now make up 31 percent.*

Vulnerabilities

First Half of 2009 Vulnerability Disclosure Count

X-Force analyzed and documented 3,240 new vulnerabilities in the first half of 2009, an 8 percent decrease in comparison to the first half of 2008 and the lowest count of new disclosures in the first half of the year in four years.

The rate of vulnerability disclosures in the past few years appears to have reached a high plateau. In 2007, the vulnerability count dropped for the first time, but then in 2008, there was a new record high. The annual disclosure rate appears to be fluctuating between 6-7 thousand new disclosures each year.

The slowing disclosure rate in the first half of this year was primarily driven by declines in some of the largest categories of vulnerabilities. Although vulnerabilities affecting Web applications continue to be the largest category of disclosure, major subcategories (SQL injection and file include) have declined, and one of the largest subcategories affecting client applications, ActiveX controls, has also declined. The slowing disclosure rate is most likely due to the disappearance of the low-hanging fruit in these highly targeted categories (for researchers and attackers alike). Unfortunately, the slowing disclosure rate is not being mirrored by attacks targeting these vulnerabilities, especially SQL injection and ActiveX controls. See Web Application Attacks on page 17 and Browser and Other Client-Side Vulnerabilities and Exploits on page 24 for more details.

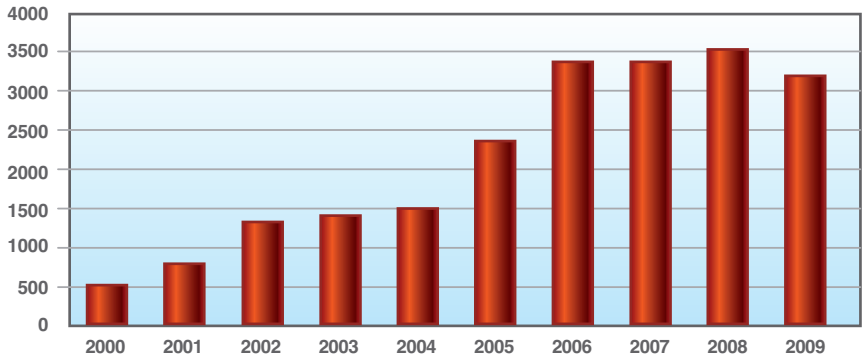


Figure 1: *Vulnerability Disclosures in the First Half of Each Year, 2000-2009*

To avoid any ambiguity regarding the characterization of vulnerabilities, the IBM Internet Security Systems (ISS) definition below is applied to this report.

Vulnerability—any computer-related vulnerability, exposure, or configuration setting that may result in a weakening or breakdown of the confidentiality, integrity, or accessibility of the computing system.

Vulnerability Disclosures by Severity

The Common Vulnerability Scoring System (CVSS) is the industry standard for rating vulnerability severity and risk based on metrics (base and temporal) and formulas. Base metrics are comprised of characteristics that generally do not change over time. Base metrics include access vector, complexity, authentication, and the impact bias. Temporal metrics are made up of characteristics of a particular vulnerability that can and often do change over time, and include the exploitability, remediation level, and report confidence.

Vulnerabilities identified as Critical by CVSS metrics are vulnerabilities that are installed by default, network-routable, do not require authentication to access and will allow an attacker to gain system or root level access.

Table 1 represents the severity level associated with the both base and temporal CVSS scores.

CVSS ScoreLevel	Severity Level
10	Critical
7.0 – 9.9	High
4.0 – 6.9	Medium
0.0 – 3.9	Low

Table 1: *CVSS Score and Corresponding Severity Level*

For more information about CVSS, a complete explanation of CVSS and its metrics are on the First.org Web site at *http://www.first.org/cvss/*.

CVSS Base Scores

As Figure 2 indicates, only about 1 percent of all vulnerabilities scored in the Critical category in the first half of 2009, similar to the percentage seen in 2008.

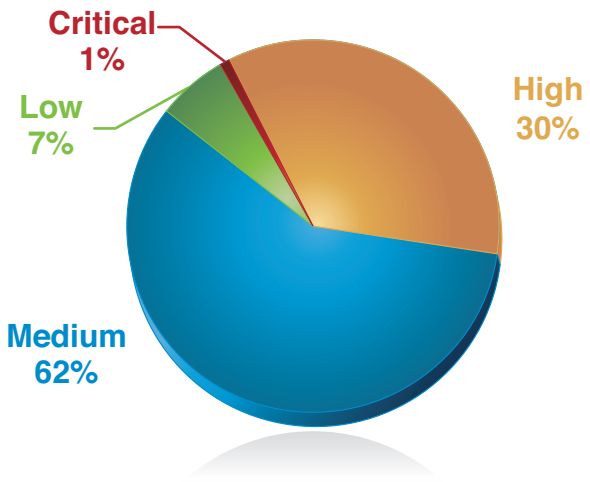


Figure 2: *CVSS Base Scores, 2009 H1*

High vulnerabilities, however, are in decline, down to 30 percent in the first half 2009 in comparison to 36 percent in 2008 as shown in Figure 3. The overall decline in the number of high severity disclosures is around 30 percent in comparison to the number disclosed in 2008. Medium severity vulnerabilities have filled the gap, comprising 62 percent of the vulnerabilities disclosed in the first half of this year up from 54 percent in 2008.

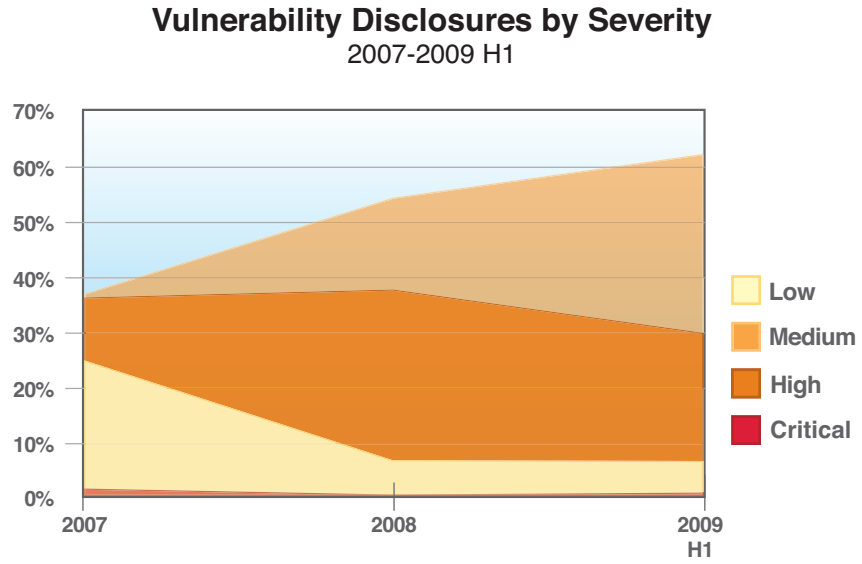


Figure 3: *CVSS Base Scores, 2007-2009 H1*

Exploitability Probability Quadrant

Although CVSS is a good mechanism for scoring the ease of exploitation and criticality of exploitation, it does not yet take into account the monetization and attacker motivation or cost to exploiting any given vulnerability. The X-Force Exploitability Probability quadrant incorporates the ease of exploitation along with the benefits and costs from the attacker perspective. Some of the most critical (and/or hyped) vulnerabilities disclosed in the first half of 2009 along with those discovered by X-Force are mapped in Figure 5. These vulnerabilities are described in detail on the X-Force Alert and Advisory page at *http://www.iss.net/threats/ThreatList.php*.

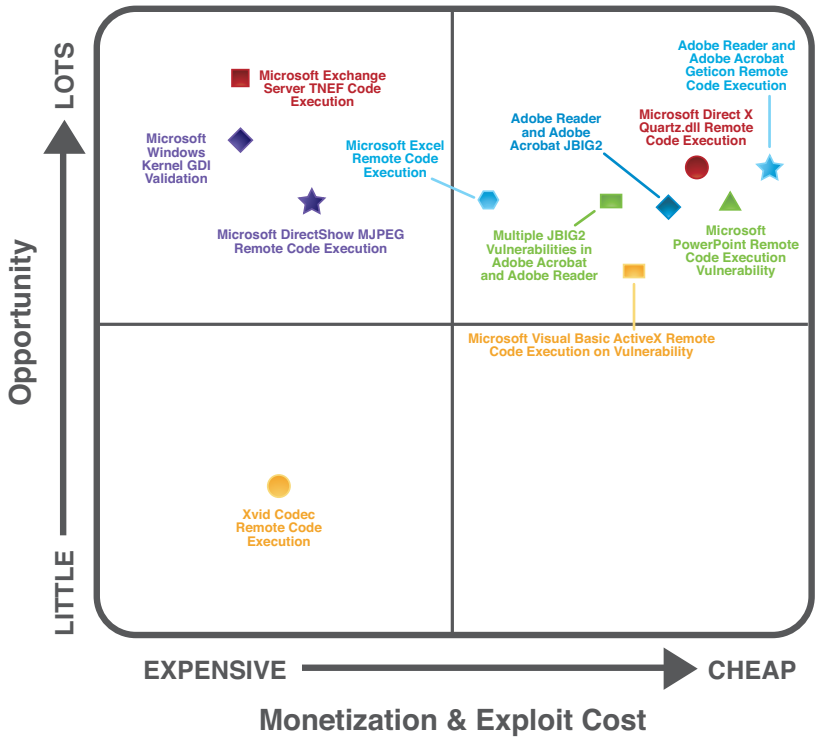


Figure 4: *X-Force Exploitability Probability Quadrant, 2009 H1*

Vendors with the Most Vulnerability Disclosures

Vulnerability disclosures for the top ten vendors in the first half of 2009 accounted for nearly a quarter of all disclosed vulnerabilities, up significantly from 2008 (5 percentage points from 19 percent) and 2007 (when they represented around 18 percent of all disclosures). Table 2 reveals who the top ten vendors are and their percentages of vulnerabilities in the first half of 2009.

These statistics do not balance vulnerability disclosures with market share, number of products, or the lines of code that each vendor produces. In general, mass-produced and highly distributed or accessible software is likely to have more vulnerability disclosures.

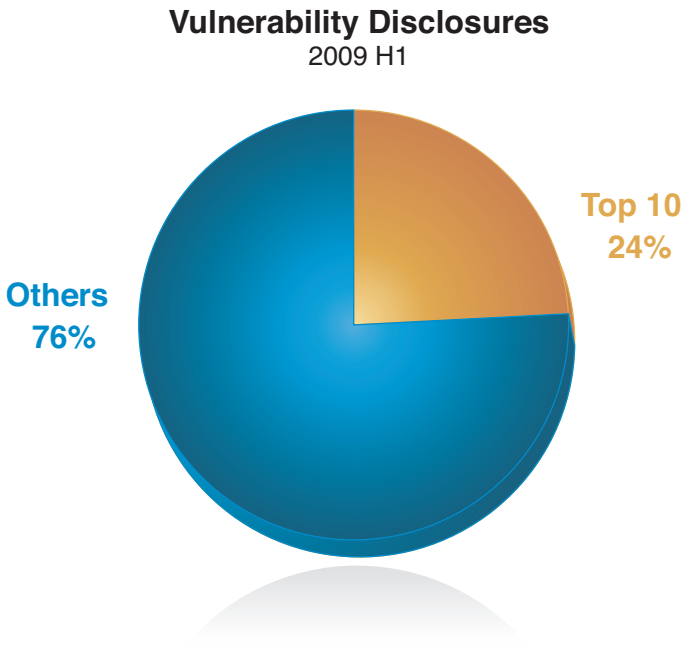


Figure 5: *Percentage of Vulnerability Disclosures Attributed to Top Ten Vendors, 2009 H1*

Major Shifts in the Top Vendor List

The X-Force database team uses an industry standard called CPE, or Common Platform Enumeration (more info at *http://cpe.mitre.org/*), to assign vulnerabilities to vendors and vendor products.

In 2008, several new vendors that produce Web application software appeared on the top ten vendor list for the first time: Joomla!, WordPress, Drupal, and TYPO3. As of the first half of 2009, only Drupal and Joomla! remained on the list. Although TYPO3 would still show up if we published a top 20, WordPress has practically dropped off the charts with less than a handful of vulnerability disclosures in the first half of this year.

Another significant change is the position of Microsoft. After holding the top vendor spot for three straight years in a row (2006/3.1 percent, 2007/3.7 percent, 2008/3.16 percent), it has dropped down to number three. Apple has taken the number one slot, and Sun, who broke the top 5 for the first time in 2008, is in second place as the vendor with the most vulnerability disclosures so far this year.

This shift for Sun is very significant, but probably not in the way you might expect. A similar jump in disclosures was seen in their Sun Solaris operating system this year—it hit the number one slot in OS with the most public vulnerability disclosures over Apple and Microsoft (see Operating Systems with the Most Vulnerability Disclosures on page 21 for details). However, if you look at the types of disclosures Sun has released, you will find that most of them appear to be internally discovered and are announced (with minimal vulnerability details—do not ask us if we are tired of writing "an unspecified vulnerability…" when it is a Sun disclosure!) along with a patch for the issue. For the vast number of disclosures Sun makes, they have an impressive patch rate (only 4 percent left unpatched). So, what these statistics really mean is that Sun has most likely implemented a more mature vulnerability discovery and reporting framework for their software. Rather than a black eye, they probably deserve a gold star for following a responsible policy of discovery and disclosure.

2009 H1			2008 (Full Year)		
Ranking	Vendor	Disclosures	Ranking	Vendor	Disclosures
1.	Apple	3.8%	1.	Microsoft	3.16%
2.	Sun	3.6%	2.	Apple	3.04%
3.	Microsoft	3.1%	3.	Sun	2.19%
4.	Oracle	2.7%	4.	Joomla!	2.07%
5.	IBM	2.5%	5.	IBM	2.00%
6.	Drupal	2.0%	6.	Oracle	1.65%
7.	Mozilla	1.8%	7.	Mozilla	1.43%
8.	Cisco	1.8%	8.	Drupal	1.42%
9.	Linux	1.5%	9.	Cisco	1.23%
10.	Joomla!	1.2%	10.	TYPO3	1.23%

Table 2: *Vendors with the Most Vulnerability Disclosures*

Availability of Vulnerability Fixes and Patches

Similar to the end of 2008, nearly half (49 percent) of all vulnerabilities disclosed in the first half of 2009 have no vendor-supplied patch at the end of the period.

The following chart provides an analysis of those vendors with twenty or more disclosures this year that have provided the fewest patches to fix those issues.

This analysis provides some interesting results. As noted in the Top Vendor section, one of the Web application vendors that appeared in the 2008 list disappeared from Top Vendor list for the first half of 2009. However, most of them reappear here (Joomla!, TYPO3, and Drupal). The second point of interest is that Sun, who took the number two spot in vendors with the most disclosures, has a very good patch rate… only missing 4 percent of all disclosures.

Certain vendors that appeared on the vendors with the most disclosures list have disappeared completely when you consider patch rate: Oracle, IBM, and Linux®.

Vendor	Disclosures	Unpatched	% Unpatched
Joomla!	40	32	80%
Apple	122	22	18%
Microsoft	100	17	17%
Drupal	65	9	14%
Mozilla	59	8	14%
TYPO3	24	3	13%
Cisco	57	5	9%
Novell	25	2	8%
HP	40	3	8%
Sun	117	5	4%

Consequences of Exploitation

X-Force categorizes vulnerabilities by the consequence of exploitation. This consequence is essentially the benefit that exploiting the vulnerability provides to the attacker. Table 3 describes each consequence.

Consequence	Definition
Bypass Security	Circumvent security restrictions such as a firewall or proxy, and IDS system or a virus scanner
Data Manipulation	Manipulate data used or stored by the host associated with the service or application
Denial of Service	Crash or disrupt a service or system to take down a network
File Manipulation	Create, delete, read, modify, or overwrite files
Gain Access	Obtain local and remote access. This also includes vulnerabilities by which an attacker can execute code or commands, because this usually allows the attacker to gain access to the system
Gain Privileges	Privileges can be gained on the local system only
Obtain Information	Obtain information such as file and path names, source code, passwords, or server configuration details
Other	Anything not covered by the other categories

Table 3: *Definitions for Vulnerability Consequences*

The most prevalent primary consequence of vulnerability exploitation continues to be Gain Access, which rebounded from a low in 2008 and, in the first half of 2009, is now re-approaching the 50 percent mark that was previously seen throughout 2006 and 2007. After a significant jump in 2008, vulnerabilities that allow an attacker to manipulate data took a plunge in the first half of this year, but are still higher in comparison to 2006 and 2007. Most of these percentage changes are due to the decline in SQL injection and file include vulnerabilities recorded for the first half of 2009.

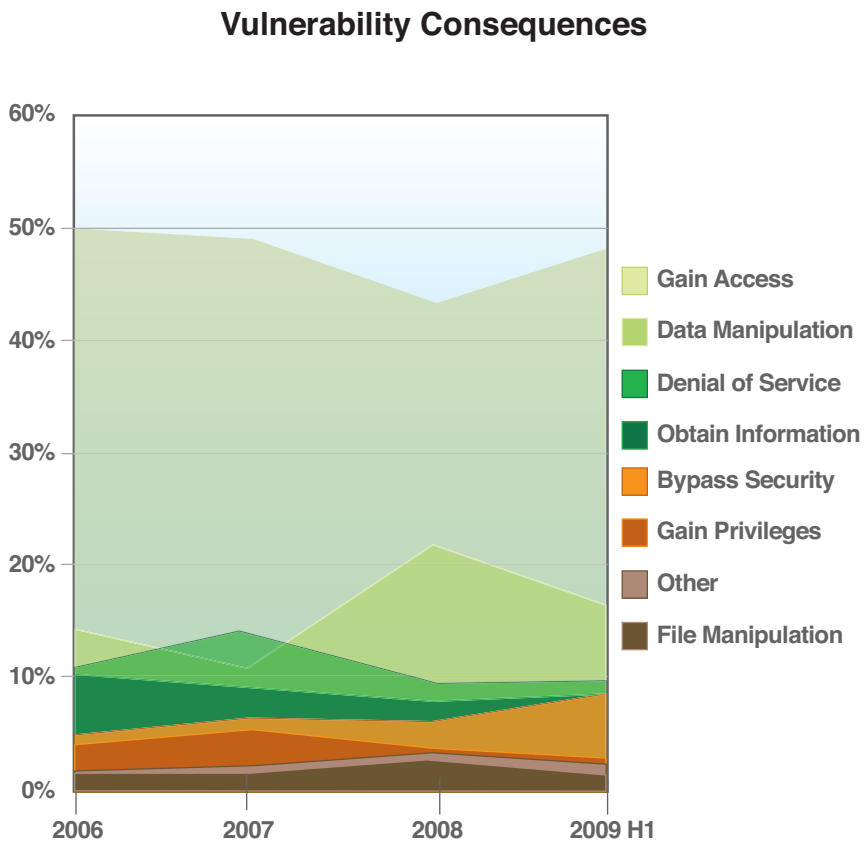


Figure 6: *Vulnerability Consequences as a Percentage of Overall Disclosures, 2006-2009 H1*

Web Application Vulnerabilities

The most prevalent type of vulnerability affecting servers today is unquestionably vulnerabilities related to Web applications.

Although the number of vulnerabilities affecting Web applications has grown at a staggering rate, the growth demonstrated in the first half of 2009 may indicate the start of a plateau, at least in standard (off-the-shelf) software applications for the Web. These figures do not include custom-developed Web applications or customized versions of these standard packages, which also introduce vulnerabilities.

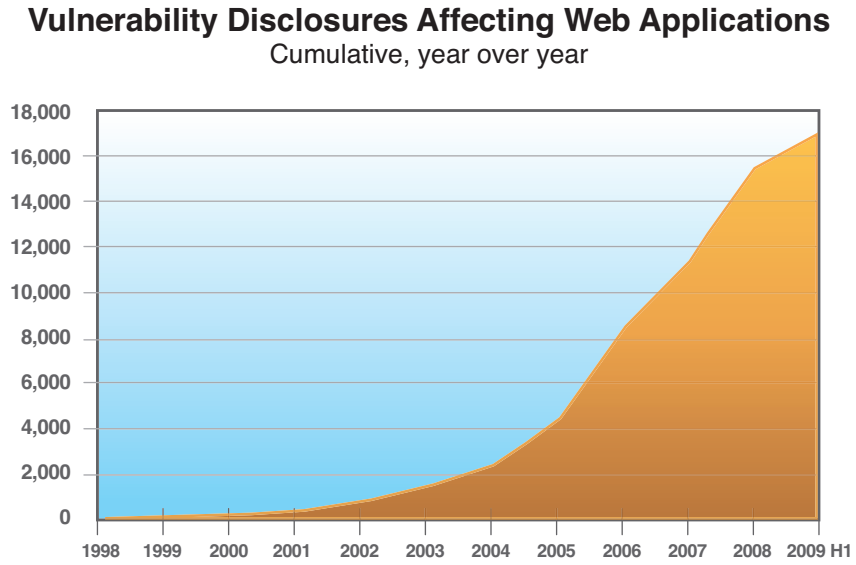


Figure 7: *Cumulative Count of Web Application Vulnerability Disclosures, 1998-2009 H1*

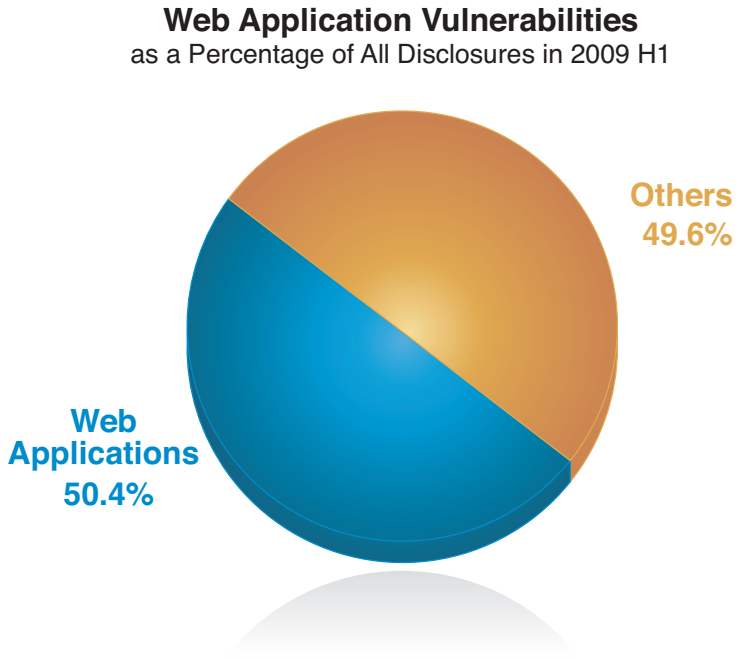


Figure 8: *Percentage of Vulnerability Disclosures that Affect Web Applications, 2009 H1*

Web Application Vulnerability Disclosures by Attack Categories

The predominate types of vulnerabilities affecting Web applications are cross-site scripting (XSS), SQL injection, and file include vulnerabilities. SQL injection and Cross-Site Scripting are neck and neck in a race for the top spot in Web application vulnerability categories. In 2008, SQL injection replaced cross-site scripting as the predominant Web application vulnerability disclosure affecting off-the-shelf Web applications, but as Figure 9 shows, researchers continue to discover many new Cross-Site Scripting issues. The spike of SQL injection vulnerabilities seen in 2008 may be attributed to the development of automated attack methods that discovered many SQL injection vulnerabilities, which were later used for attacks, on live Web sites. For many security administrators and researchers, these automated tools put increased pressure on them to find SQL injection vulnerabilities before the attackers did.

Figure 9 shows how SQL injection and other major categories of Web application vulnerabilities have changed over the years, and Table 4 describes each category including the impact they can have on organizations and the customers they serve.

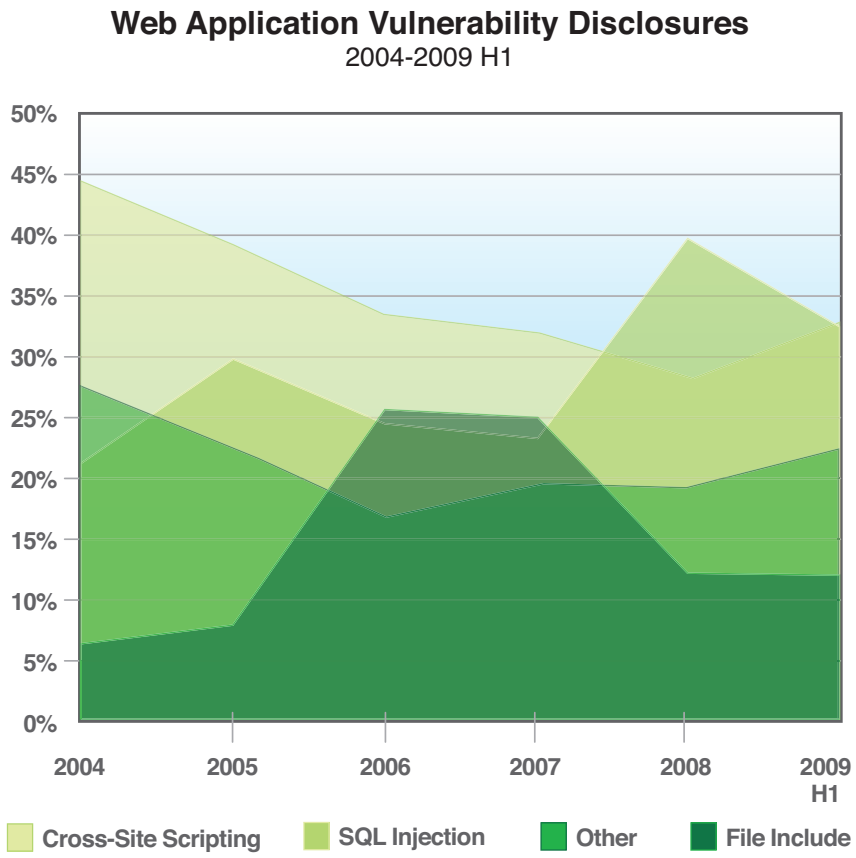


Figure 9: *Web Application Vulnerabilities by Attack Technique, 2004-2009 H1*

Attack Technique	Description
Cross-site Scripting	Cross-site scripting vulnerabilities occur when Web applications do not properly validate user input from form fields, the syntax of URLs, etc. These vulnerabilities allow attackers to embed their own script into a page the user is visiting, manipulating the behavior or appearance of the page. These page changes can be used to steal sensitive information, manipulate the Web application in a malicious way, or embed more content on the page that exploits other vulnerabilities. The attacker first has to create a specially-crafted Web link, and then entice the victim into clicking it (through spam, user forums, etc.) The user is more likely to be tricked clicking the link, because the domain name of the URL is a trusted or familiar company. The attack attempt may appear to the user to come from the trusted organization itself, and not the attacker that compromised the organization's vulnerability.
SQL Injection	SQL injection vulnerabilities are also related to improper validation of user input, and they occur when this input (from a form field, for example), is allowed to dynamically include SQL statements that are then executed by a database. Access to a back-end database may allow attackers to read, delete, and modify sensitive information, and in some cases execute arbitrary code. In addition to exposing confidential customer information (like credit card data), SQL injection vulnerabilities can also allow attackers to embed other attacks inside the database that can then be used against visitors to the Web site.
File Include	File include vulnerabilities (typically found in PHP applications) occur when the application retrieves code from a remote source to be executed in the local application. Oftentimes, the remote source is not validated for authenticity, which allows an attacker to use the Web application to remotely execute malicious code.
Other	This category includes some denial-of-service attacks and miscellaneous techniques that allow attackers to view or obtain unauthorized information, change files, directories, user information or other components of Web applications.

Table 4: *Description of the Most Prevalent Categories of Web Application Vulnerabilities*

Web Application Attacks

The IBM ISS Managed Security Service (MSS) data also provides real-world insight into the most prevalent types of Web application vulnerabilities and their exploitation. Similar to vulnerability disclosures, cross-site scripting and injection attacks dominate the attack landscape.

Cross-Site Scripting Attacks

60 percent of the cross-site scripting attacks involve the use of a <SCRIPT> tag in URL or CGI data, which can indicate an attack attempt against the Web server.

Injection Attacks

The vast majority of injection attacks are attributed to SQL-related attacks (about 90 percent). 70 percent of SQL-related attacks are SQL injection. The second most prevalent type of SQL attack is related to select statements, which are typically attempts at retrieving sensitive data stored in the back-end database.

Information Disclosure Attacks

The third-largest category of detected attacks is the Information Disclosure category, and the most prevalent attack, representing 70 percent, is an attempt to grab the Unix password file (the "passwd" or "shadow" password file) on the server using an HTTP GET request.

Web Application Attack Chart

The following chart provides an overview of the most prevalent types of Web application exploits as seen in our global MSS operations, and the table below it provides a definition for the attack categories. Unfortunately, many Web sites incorporate code that introduces vulnerabilities to support a feature or function, such as using SQL injection to get data from a Web form, so some legitimate usage may look like an attack attempt.

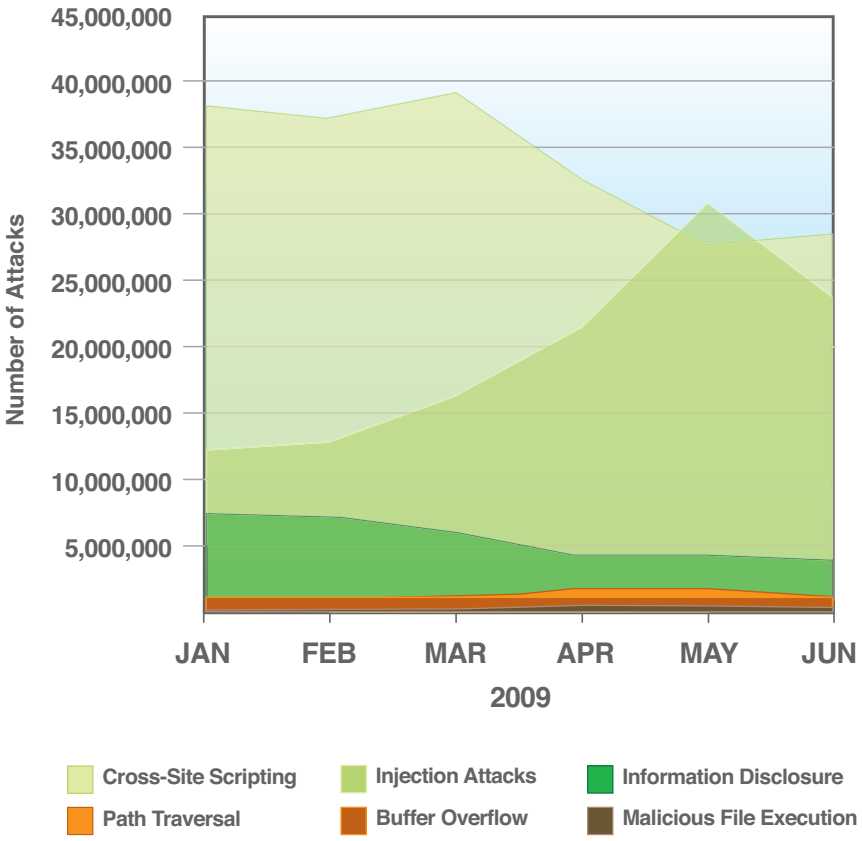


Figure 10: *Web Application Attacks by Category, IBM ISS Managed Security Services 2009 H1*

Attack Category	Description
Buffer Overflow attacks	This type of attack overflows a buffer with excessive data, which allows an attacker to run remote shell on the computer and gain the same system privileges granted to the application being attacked.
Cross-site Scripting attacks	This type of attack exploits the trust relationship between a user and the Web sites they visit.
Information Disclosure attacks	This type of attack is aimed at acquiring system specific information about a Web site including software distribution, version numbers, and patch levels. The acquired information might also contain the location of backup files or temporary files.
Injection attacks	This type of attack allows an attacker to inject code into a program or query or inject malware onto a computer in order to execute remote commands that can read or modify a database, or change data on a Web site.
Malicious File Execution attacks (also known as file include attacks)	This type of attack allows an attacker to perform remote code execution, remote root kit installation, complete system compromise, and internal system compromise (on Windows systems) through the use of SMB file wrappers for the PHP scripting language.
Path Traversal attacks	This type of attack forces access to files, directories, and commands that are located outside the Web document root directory or CGI root directory.

***Table 5**: Description of the Most Prevalent Categories of Web Application Attacks*

Automated SQL Injection Probes and Attacks

In 2008, SQL injection hit a high point not only in terms of vulnerability disclosures, but also in terms of exploitation. Automated toolkits appeared on the threat scene in the summer of 2008 and have continued to flourish in 2009. Botnets like ASPROX and NV32ts have incorporated tactics that seek out vulnerable Web sites and report their finds back to the botnet operators. From there, attackers can devise specific attacks to make full use of the vulnerabilities and data available for compromise. The volume of SQL injection attacks continues to increase in 2009. Although the most dramatic rise occurred between mid 2008 and the end of 2008 (30x increase in attacks), the SQL injection attack volume has continued to grow since the end of 2008 growing 50 percent in Q1 in comparison to 2008 Q4 and nearly doubling in Q2 (in comparison to Q1). Month over month growth is more sporadic, but peak growth months were April, with 46 percent growth over March, and May with a record high growth of 76 percent in comparison to April.

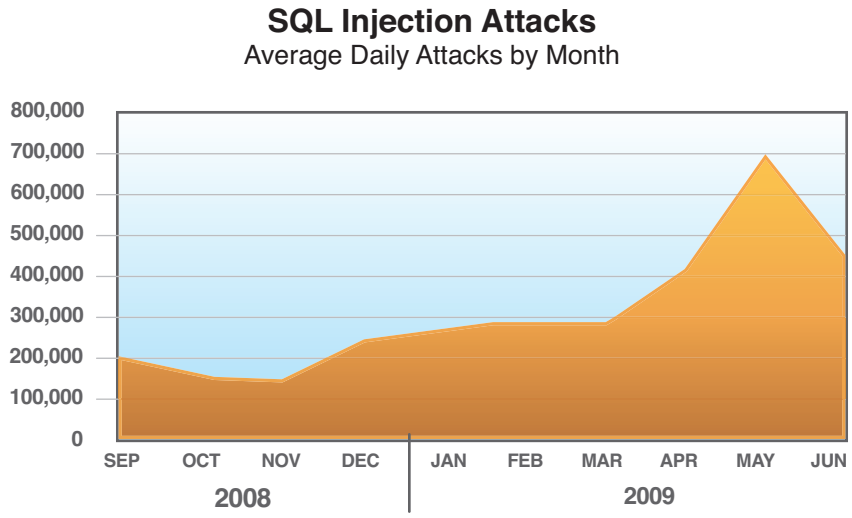


Figure 11: *SQL Injection Attacks Monitored by IBM ISS Managed Security Services, Q4 2008–2009 H1*

Operating Systems with the Most Vulnerability Disclosures

In the 2008 report, X-Force presented an analysis of operating systems with the most vulnerabilities. These vulnerabilities were counted according to how each vendor reports their platforms through the Common Platform Enumeration (or CPE). There are slight differences in how some vendors classify their platforms. For example, Linux has a platform called "Linux kernel," but vulnerabilities reported for that "platform" may also affect other Linux versions even though they may not be officially reported for that platform as it is reported in CPE. Other differences included the way that vendors classify a platform. Apple, for example, combines all versions of their Apple Mac OS X software into a single "platform" and only differentiates between the server and desktop versions of the software. Microsoft calls each of its major operating systems "platforms" even though some of these platforms may be considered by other individuals to be "versions" of Windows.

So, instead of counting vulnerabilities according to the named "platforms" in CPE, here is a slightly different analysis that counts each unique vulnerability reported for a genre of operating systems. For example, this analysis compares all vulnerabilities reported for Microsoft operating systems and compares them to all of the vulnerabilities reported for Apple operating systems in any given year. If a certain vulnerability applies to multiple versions of operating systems in that genre, it is only counted one time. For example, if a certain CVE applies to both Apple Mac OS X and also Apple Mac OS X Server, it is only counted one time for the Apple genre.

The results are not entirely dissimilar from the 2008 analysis. Apple would still have been in the top slot for 2008, and, if it had not have been for the sudden rise in Sun disclosures (see Major Shifts in the Top Vendor List on page 9 for more details), Apple still would have been in the lead with Linux closely behind it, and Microsoft would still take fourth place behind Sun Solaris (third).

For the first half of this year, Sun Solaris has leapt to the top (as described in the Top Vendor section, mostly likely due to a positive change in their vulnerability disclosure policy). The remaining operating systems are holding steady, keeping the same relative position as in 2008, after accounting for the shift in Sun Solaris, with one exception. BSD is now in the number five slot, replacing IBM AIX who was fifth in 2008.

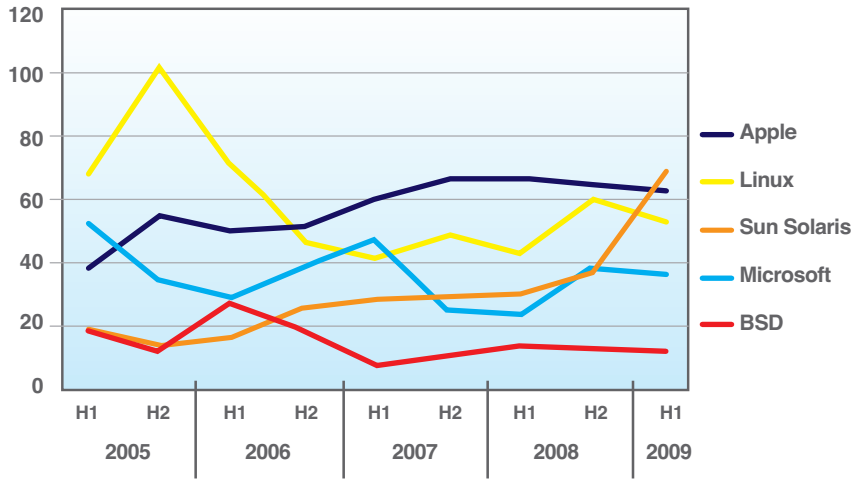


Figure 12: *Vulnerability Disclosures Affecting Operating Systems, 2005-2009 H1*

Focusing on critical and high vulnerabilities is another way to look at this issue. From a protection standpoint, these high-severity vulnerabilities are typically the ones we most worry about since they often lead to complete remote compromise, the prize possession of attackers. When you filter out the mediums and lows, Microsoft operating systems take first place in 2008 and the first half of 2009, although they have dropped significantly over the past six months on average. Apple, Sun Solaris, and Linux are in a close race for second, third, and fourth place, while IBM AIX does show up, again, here in fifth place.

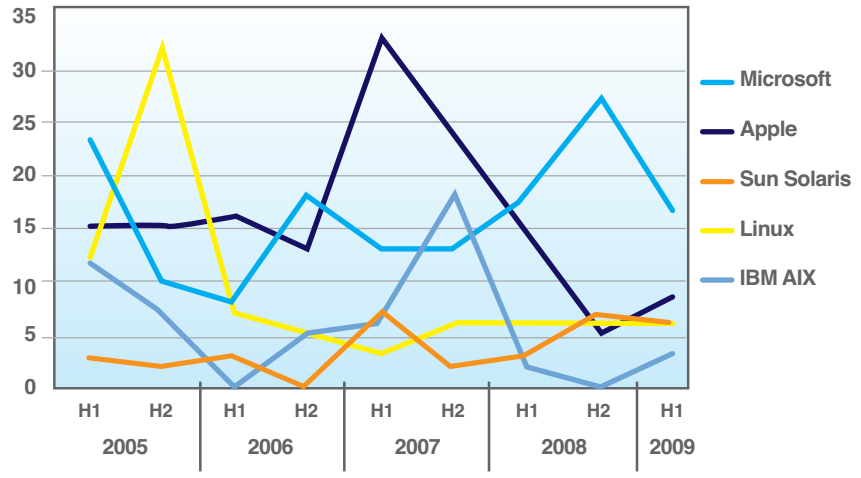


Figure 13: *Critical and High Vulnerability Disclosures Affecting Operating Systems, 2005-2009 H1*

Using this new methodology, the top operating systems in each category account for 89 percent of all operating system vulnerability disclosures and 93 percent of all critical and high operating system disclosures in the first half of 2009. Details are in the following chart:

Operating System	Percentage of Critical and High	Percentage of all OS Vulnerabilities
Microsoft	39%	14%
Apple	18%	24%
Sun Solaris	14%	26%
Linux	14%	20%
IBM AIX	7%	3%
BSD	2%	4%
Others	7%	11%

Table 6: *Operating Systems with the Most Critical and High Vulnerability Disclosures, 2009 H1*

Browser and Other Client-Side Vulnerabilities and Exploits

Vulnerabilities affecting personal computers are the second-largest category of vulnerability disclosures after Web application vulnerabilities and represent around one fifth of all vulnerability disclosures.

Client-side vulnerabilities: Vulnerabilities affecting the operating system or applications running on personal computers. In addition to the core operating system, vulnerable components could include e-mail clients, Web browsers, document viewers, and multimedia applications.

Client-Side Vulnerabilities—Document Format Vulnerabilities Increasing

Following the trend noted in our 2008 report, browser and operating system vulnerabilities continue to decline. Rapidly taking up the slack are vulnerabilities found in document and multimedia applications. Figure 14 shows the changes in critical and high vulnerability disclosures for the top categories of vulnerabilities affecting personal computers in 2009, and how they have changed since 2005. For a comparison of which vulnerabilities are most often exploited, see Exploitation Trends on page 27.

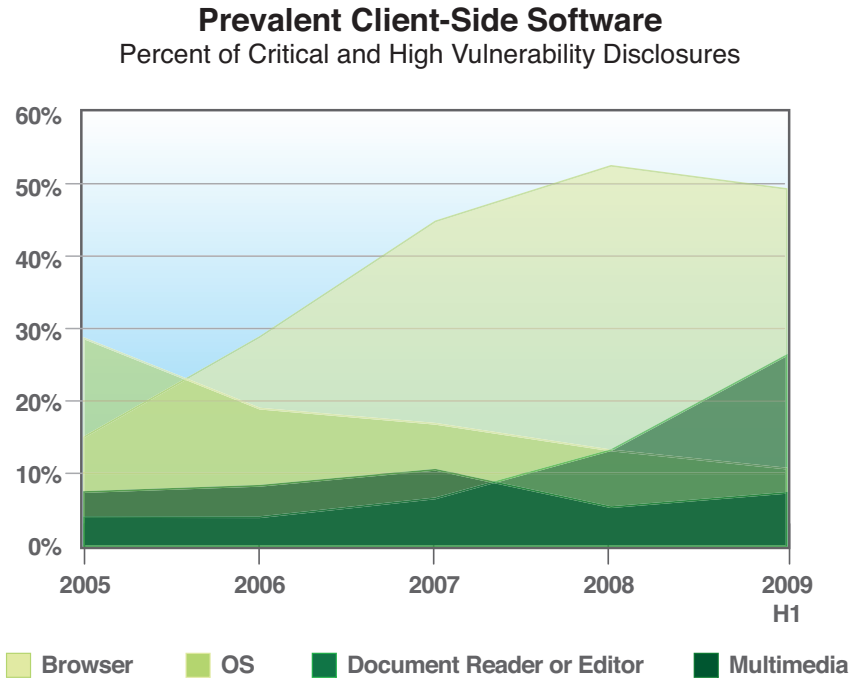


Figure 14: *Critical and High Vulnerability Disclosures Affecting Client-Side Applications by Application Category, 2005-2009 H1*

Document Format Vulnerabilities

Document format vulnerabilities affect more than simply the document reader itself. Vulnerabilities can be related to browser plug-ins for that file type or even servers that process files as they are sent to or requested by end users. In the past, most of these vulnerabilities were related to familiar Office document formats, such as .doc, .xls, .ppt, etc. However, in 2009, the rate of disclosures related to Portable Document Format (PDF) vulnerabilities skyrocketed, and the number disclosed in the first half of 2009 alone has already surpassed the number of disclosures that occurred over the full year of 2008 and also traded places with Office document disclosures, and is now the number one type of document-related vulnerability disclosure.

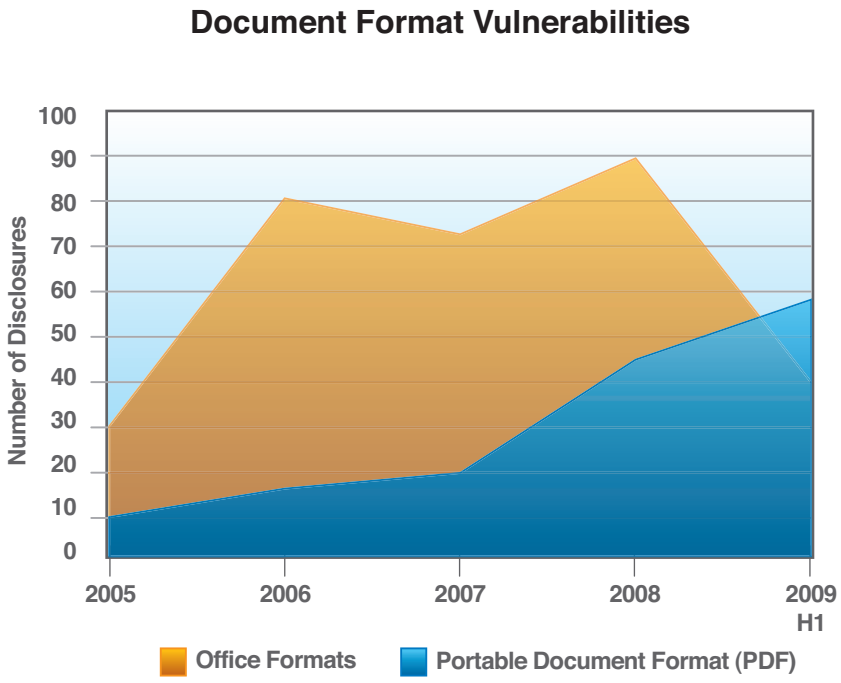


Figure 15: *Vulnerability Disclosures Related to Document Format Issues, 2005-2009 H1*

Browser Vulnerabilities—Firefox Surpasses Internet Explorer

Even with the rise in document format vulnerabilities, the largest number of client-side vulnerabilities released in the first half of 2009 affects Web browsers and their plug-ins. The most affected component out of all the browsers and types of plug-ins is the ever-pervasive ActiveX control, although the rate of disclosures for ActiveX controls is continuing to slow as shown in Figure 16.

After a record low in 2007, the Mozilla Firefox Web browser continues to take an increasing percentage of Web browser vulnerabilities, and actually surpassed the number of disclosures for Microsoft Internet Explorer during the first half of this year.

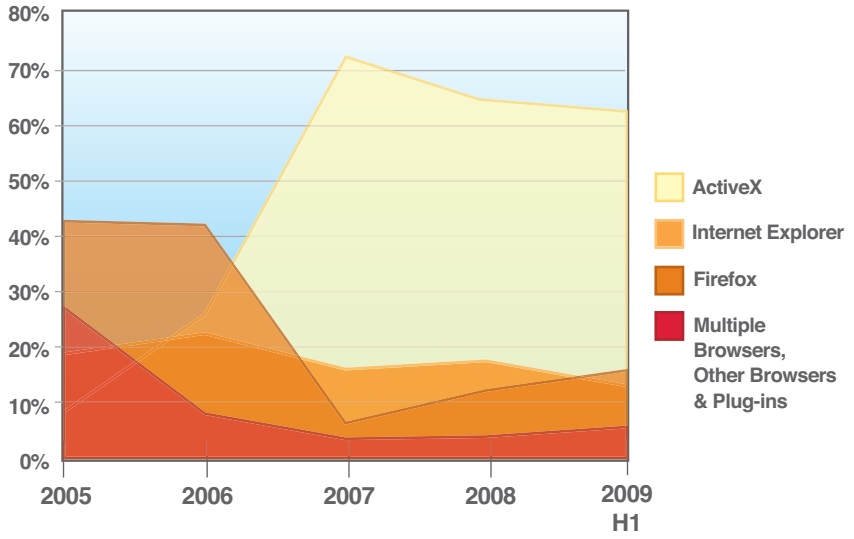


Figure 16: *Critical and High Vulnerability Disclosures Affecting Browser-Related Software, 2005-2009 H1*

Unfortunately, the decline in ActiveX disclosures does not appear to be making an impact on exploitation. As with other browser-related vulnerabilities, many attackers rely upon users who do not keep their browsers current. Although Microsoft has made great strides in preventing ActiveX exploitation through changes to Microsoft Internet Explorer, exploitation remains an issue, and attackers are discovering more and more ways to exploit ActiveX and other browser vulnerabilities either before or shortly after vulnerability disclosure and the availability of patches.

Exploitation Trends

X-Force continues to track growth in Web browser exploitation through its Whiro crawlers, which combine independent analysis with IBM ISS Managed Security Services operational alerting data. X-Force has developed specialized technology to identify exploits used even in the most obfuscated cases including where toolkits attempt multiple exploits.

During 2008, it became clear that lone Web browser exploits in the wild were dying out and being replaced by the organized use of Web exploit toolkits. These toolkits can deliver all of the exploits at once to Web site visitors, or the toolkit can select specific exploits based on data, such as:

- *Browser cookie set by the toolkit*
- *Browser agent used by the victim*
- *Geographic location derived from the victim's IP address*
- *Referrer URL (the URL that directed the victim to the Web site)*

Some more recent tricks include making multiple queries to dynamically-generated URLs that only host one or two exploits per URL. In this way, if you discover one page to block, another page could sneak past if undetected. Similar techniques involve using staged exploit code that dynamically requests additional code from a remote server to deliver one or more exploits. In this case, it is a means to better-obfuscate the code, attempt to evade detection, and potentially foil blacklisting with dynamic paths.

Deployments of exploit toolkits may be organized by a group or are, in some cases, financially supported by multiple attackers who are credited by an ID number associated in their attack URLs. The use of IDs by multiple attackers is interesting because it allows attackers to get a piece of the action with a smaller initial investment. Of further interest is the rampant copying of exploit code in exploit toolkits. Although it's unknown how many attackers pay for their toolkits, it would be surprising if any of them did unless the toolkit was frequently updated with new exploits.

Most Popular Exploits

Rank	2008 H2	2009 H1
1.	Microsoft MDAC RDS Dataspace ActiveX (CVE-2006-0003)	Microsoft MDAC RDS Dataspace ActiveX (CVE-2006-0003)
2.	Microsoft WebViewFolderIcon ActiveX (CVE-2006-3730)	Microsoft Snapshot Viewer ActiveX (CVE-2008-2463)
3.	Internet Explorer "createControlRange" DHTML (CVE-2005-0055)	Adobe Acrobat and Reader Collab. CollectEmailInfo (CVE-2007-5659)
4.	RealPlayer IERPCtl ActiveX (CVE-2007-5601)	Microsoft IE7 DHTML Object Reuse (CVE-2009-0075)
5.	Apple QuickTime RSTP URL (CVE-2007-0015)	RealPlayer IERPCtl ActiveX (CVE-2007-5601)

Table 7: *Most Popular Web Browser Exploits, 2008 H2 – 2009 H1*

Compared with our 2008 report, there are only two remaining exploits from the previous list: MDAC and RealPlayer IERPCtl. Sustaining only two exploits over a period of six months is much lower than the four out of five that remained from 1H 2008 to FY 2008 and indicative that the trend of exploits lasting a long time on the top five list might be changing. Interestingly, a vulnerability current to the 1st half of 2009 (CVE-2009-0075) made the top five list, which is usually populated exclusively by older vulnerabilities.

Three of the five most popular exploits are ActiveX controls. The trend of using ActiveX controls for exploitation in general still remains strong through the first half of 2009, as the following chart shows from IBM ISS Managed Security Services:

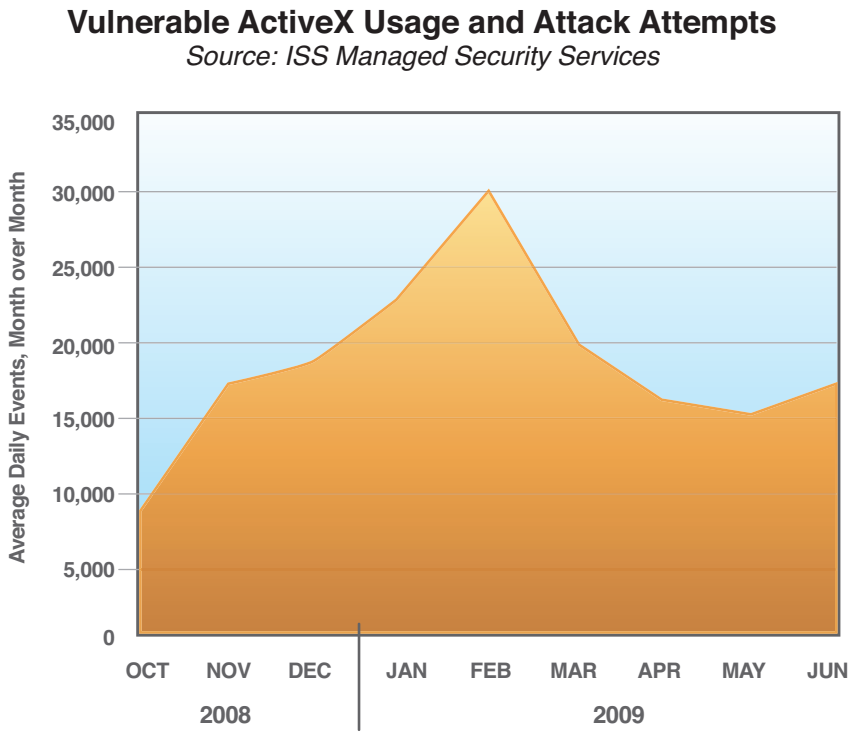


Figure 17: *Vulnerable ActiveX Control Usage and Exploitation*

The other big news is that this is the first time that a PDF exploit has pierced the top five list. While our Web exploit crawler did see some PDF getIcon exploits (CVE-2009-0927) and a significant amount of PDF util.printf() (CVE-2008-2992) exploits in the wild, neither were enough to reach the top five list this time around. Looking into the future, X-Force suspects that PDF will have at least one top five entry for the full-year 2009 report.

Most Popular Exploit Toolkits (2008 H2-2009 H1)

Rank	2008 H2	2009 H1
1.	CuteQQ	CuteQQ
2.	AdM	Tornado + IcePack Platinum
3.	mPack (and variants)	Unknown1
4.	Neosploit	Unknown2
5.	Tornado (and variants)	LuckySploit

Table 8: *Most Popular Exploit Toolkits, 2008 H2 – 2009 H1*

It is becoming increasingly difficult to identify specific exploit toolkits in the wild, because exploit toolkit code is so frequently copied. As a result, we may report a toolkit as unique, as a variant, or as both unique and as a variant. For example, the IcePack Platinum edition is effectively a fork of Tornado and the heuristics we use to identify these two are identical, so they share an entry. To reiterate, the IcePack Platinum edition code is completely different than the legacy IcePack code. Another recent example from 2008 is the CuteQQ pack which is derived from SmartPack, which itself is a variant of FirePack (and perhaps other stuff).

Currently, two of the top five most popular exploit toolkits lack specific names or even indications of being a variant of a named toolkit, which illustrates the ever-increasing amount of toolkits in the wild requiring ongoing efforts to monitor them properly. As time goes by, most of the unnamed toolkits will become known by name. By taking a sampling of crawling data which fits the profile of what a toolkit would look like, but for which no heuristic match was made, we create new heuristics to track these kits. In our year-end report for 2008, we suggested that this could become a "lesser trend" when we learn toolkit names at a later point in time, but now it looks like it is to become a growing trend.

Obfuscation

As noted in our latest report, X-Force observed a reduction in obfuscation during the second half of 2008. Specifically, we noticed a reduction in the use of using multiple layers of obfuscation and self-decoding code. While obfuscation continues to stay in flux, an increase in obfuscation intensity was observed during the 1st half of 2009.

The level of obfuscation found in Web exploits, and, especially, PDF files continues to rise, and some of these techniques are being passed to malicious multimedia files as well. From Q1 to Q2 alone, the amount of suspicious, obfuscated content monitored by IBM ISS Managed Security Services nearly doubled.

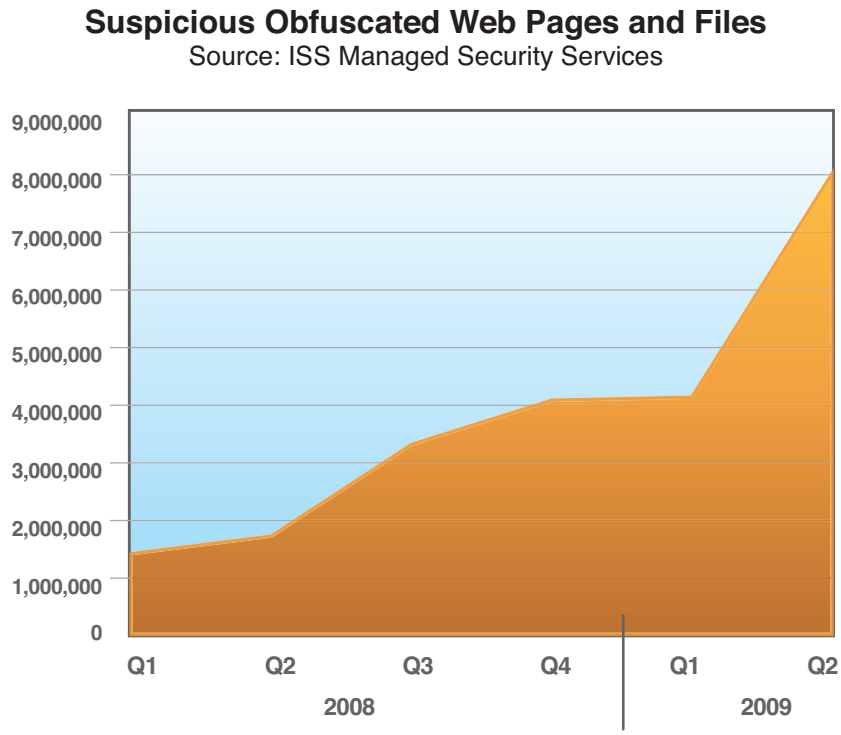


Figure 18: *Obfuscated Web Pages and Other Files, 2008-2009 H1*

On a recalculated basis, Visual Basic Script (VBScript) utilization during 2008 reached 13 percent of malicious sites. Since 2008, the prevalence of VBScript has continued to increase. Using new metrics, the number reaches 20 percent of malicious sites and an astonishing 36 percent of malicious sites when including sub-domains. To explain the 16 percent gap with the inclusion of sub-domains, consider that some sites, for example blogs, use subdomains instead of unique paths for content management. Still, an increase from 13 percent to 20 percent is quite significant. Interestingly, the VBScript observed in these browser attacks is, at times, only used for obfuscation and the final attack code may be in JavaScript. X-Force thinks the use of VBScript will continue to increase over time, although the speed of adoption might decrease for a while.

We continue to observe string replacements using regular expressions to clean up heavy obfuscation as well as string concatenations post a decoder stub such as base64 decoding. X-Force will continue to monitor whatever the state-of-the-art brings in terms of Web browser exploit obfuscation.

Web Content Trends

This section summarizes the amount and distribution of "bad" Web content that is typically unwanted by businesses based on social principles and corporate policy. Unwanted or "bad" Internet content is associated with three types of Web sites: adult, social deviance and criminal. Table 9 lists the IBM ISS Web filter categories that correspond with these types of sites.

The Web filter categories are defined in detail at:

http://www.ibm.com/services/us/index.wss/detail/iss/a1029077?cntxt=a1027244

Web Site Type	Description & Web Filter Category
Adult	Pornography Erotic / Sex
Social Deviance	Political Extreme / Hate / Discrimination Sects
Criminal	Anonymous Proxies Computer Crime / Hacking Illegal Activities Illegal Drugs Malware Violence / Extreme Warez / Software Piracy

Table 9: *Web Filter Categories Associated with Unwanted Web Content*

This section provides analysis for:

- *Percent and distribution of Web content that is considered bad, unwanted, or undesirable*
- *Increase in the amount of anonymous proxies*
- *Malware URLs: Hosting Countries and Linkage*

Analysis Methodology

X-Force captured information about the content distribution on the Internet by counting the hosts categorized in the IBM ISS Web filter database. Counting hosts is an accepted method for determining content distribution and provides the most realistic assessment. When using other methodologies – like counting Web pages/sub pages – results may differ. The IBM ISS data center is constantly reviewing and analyzing new Web content data. Consider the following statistics related to the IBM ISS data center:

- *Analyzes 150 million new Web pages and images each month*
- *Has analyzed 10 billion Web pages and images since 1999*

The IBM ISS Web Filter Database has:

- *68 filter categories*
- *105 million entries*
- *150,000 new or updated entries added each day*

Percentage of Unwanted Internet Content

As Figure 19 shows, about 8 percent of the Internet currently contains unwanted content such as pornographic or criminal Web sites.

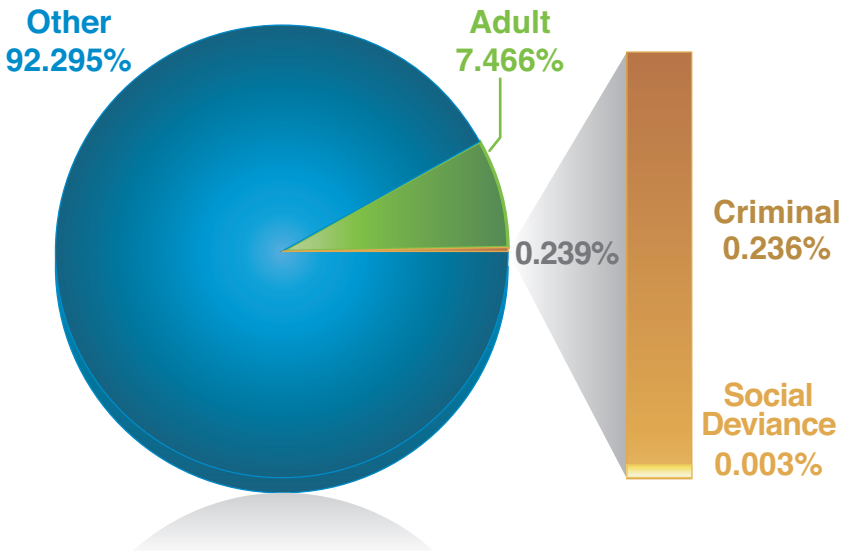


Figure 19: *Content Distribution of the Internet, 2009 H1*

Increase of Anonymous Proxies

As the Internet becomes a more integrated part of our lives not only at home, but also at work and at school, organizations responsible for maintaining acceptable environments are increasingly finding the need to put controls on where people can browse in these public settings.

One such control is a content filtering system that prevents access to unacceptable or inappropriate Web sites as described in this section of the Trend Report. In an effort to circumvent Web filtering technologies, some individuals might attempt to use an anonymous proxy (also known as Web proxy).

Web proxies allow users to enter an URL on a Web form instead of directly visiting the target Web site. Using the proxy hides the target URL from a Web filter. If the Web filter is not also set up to monitor or block anonymous proxies, then this activity, which would have normally been stopped, will bypass the filter and allow the user to reach the disallowed Webpage.

The volume of anonymous proxy Web sites reflects this trend:

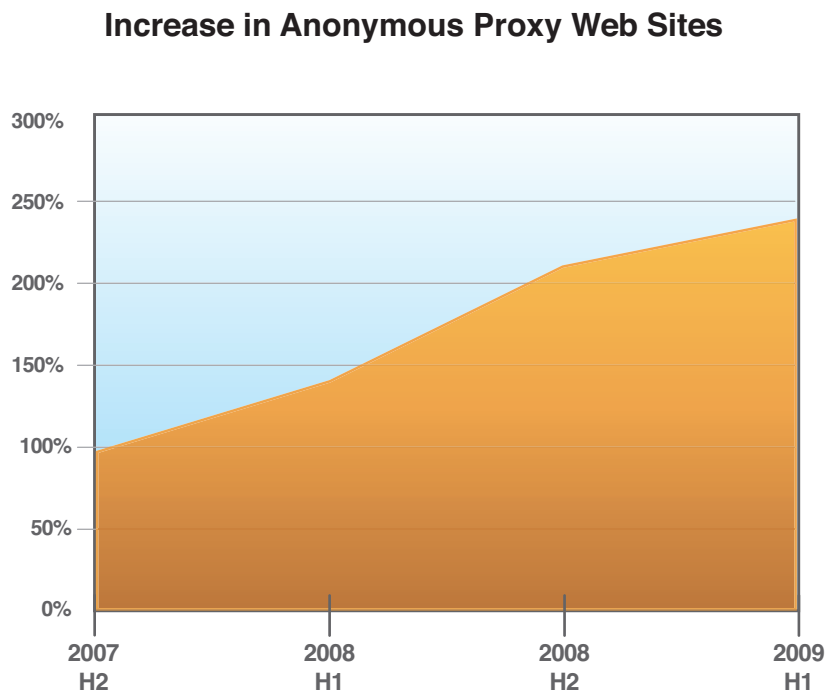


Figure 20: *Volume of Anonymous Proxy Web Sites, 2007 H2–2009 H1*

Although the increase flattens a bit, there are now considerably more than twice as many anonymous proxy Web sites online than 18 months ago.

Anonymous proxies are an incredibly important type of Web site to track, because of the growth and the ease at which they allow people to hide potentially malicious intent. The following data provides an analysis of the sites and where they are hosted.

Top Level Domains of Anonymous Proxies

The first chart shows the Top Level Domains of the newly registered anonymous proxies.

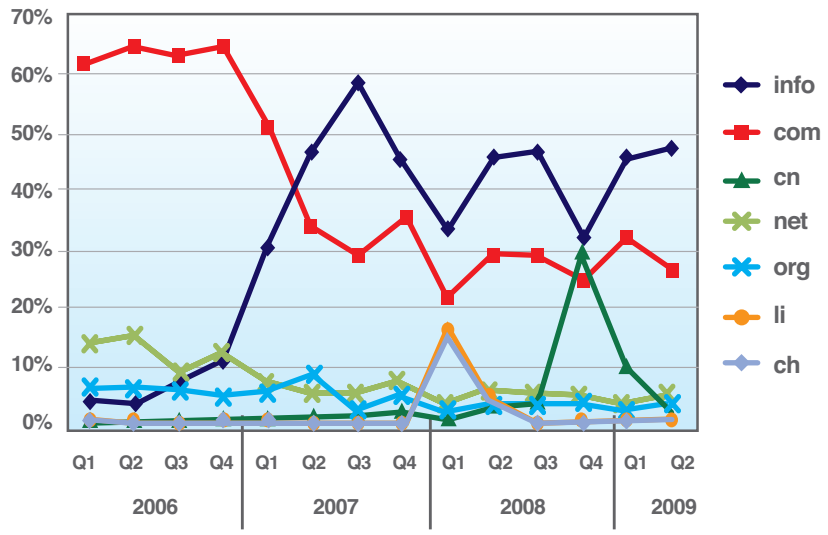


Figure 21: *Top Level Domains of Newly Registered Anonymous Proxy Web Sites, 2006-2009 H1*

In 2006, more than 60 percent of all newly registered anonymous proxies were .com domains, but since the middle of 2007, .info has been at the top. However, there were some brief interludes when other Top Level Domains became popular, like at the beginning of 2008 where the Top Level Domains of neighboring countries Switzerland and Liechtenstein together reached about 30 percent of the newly registered anonymous proxies. In the fourth quarter of 2008, the Top Level Domain of China reached nearly 30 percent of the newly registered anonymous proxies.

In any case, it is curious that .info is the predominant anonymous proxy domain. A reason could be that .com is running out of names. In the past, anonymous proxy Web sites were named something like proxy4u.info or unblockit.info and so on. In the meantime, names are chosen that do not appear to be a proxy like anyword.info, for example. Independent from using "prox" in the name or not, within .com, most domains like anyword.com are already registered (in many cases they are parked). Thus, it is much easier to register a catchy domain in the .info Top Level Domain.

Country Hosts of Anonymous Proxy Web Sites

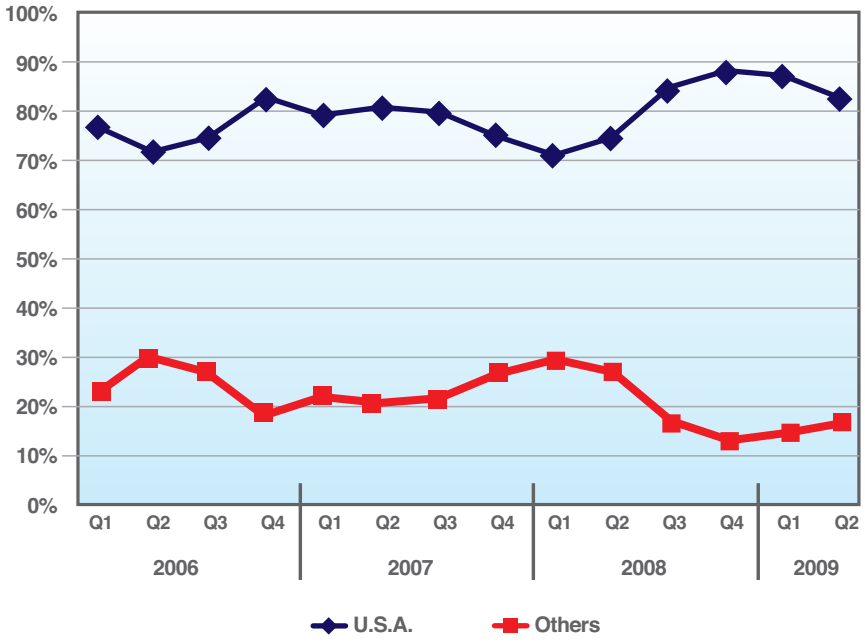


Figure 22: *Countries Where Newly-Registered Anonymous Proxy Web Sites are Hosted – United States Versus Other Countries, 2006-2009 H1*

For anonymous proxy hosting countries, the United States has held the top position for years—more than 70 percent of all newly-registered anonymous proxies have been hosted in the US over the past three and a half years. In the past 12 months, their share has climbed to more than 80 percent. All other countries host less than 10 percent of anonymous proxies, with the exception of Canada, which hosted 16.2 percent of all newly-registered anonymous proxies at the beginning of 2008.

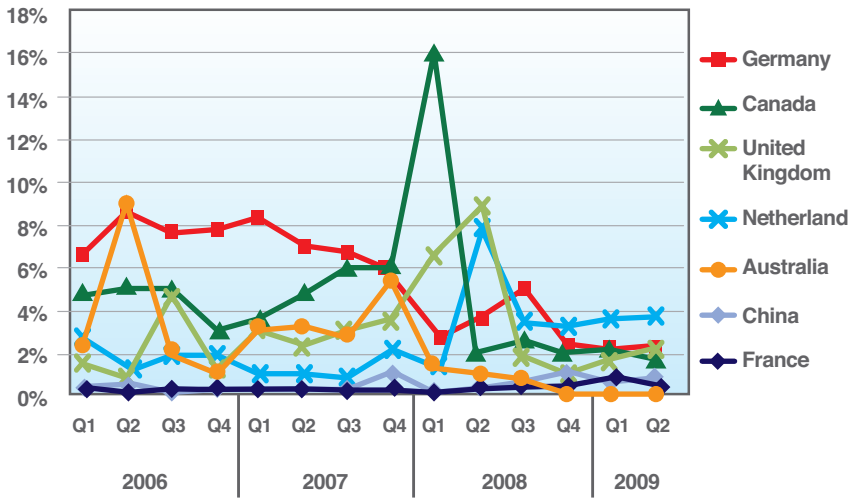


Figure 23: *Countries Where Newly-Registered Anonymous Proxy Web Sites Are Hosted –Other Countries, 2006-2009 H1*

Malicious Web Sites

The number of new malicious Web links discovered in the first half of 2009 increased by 508 percent in comparison to the number discovered in the first half of 2008. Exploitation Trends on page 27 in the Browser and Other Client-Side Vulnerabilities and Exploits section talks about the Web exploit toolkits involved in the majority of these malicious Web sites. This section discusses the countries responsible for hosting the majority of the malicious links along with the types of Web sites that most often link back to these malicious Web sites.

Geographical Location of Malicious Web Links

The United States and China continue to reign as the top hosting countries for malicious links. Although China surpassed the US for the first time at the end of 2008, the US has regained its territory and, for the first half of 2009, is the top hoster claiming 36 percent of all malicious Web links. Japan, surprisingly, after being nearly off the malicious Web link radar, is in third place, claiming 8 percent of all malicious links for the first half of 2009 as shown in Figure 24.

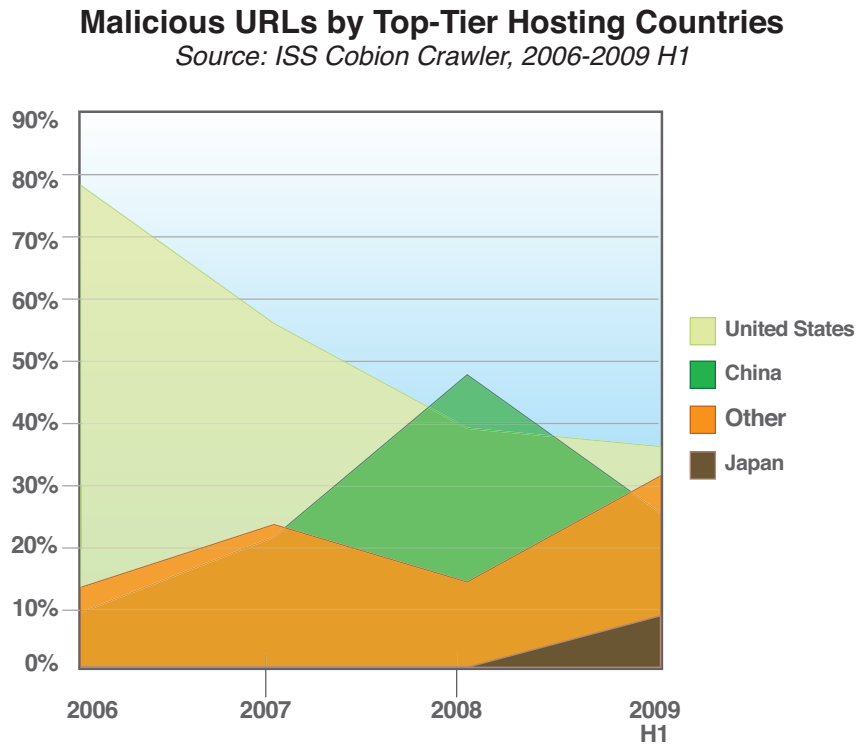


Figure 24: *Countries Hosting the Most Malicious URLs, 2006-2009 H1*

The second tier of countries (those hosting 2 to 4 percent of links) have also shifted, and, most significantly, many more countries seem to be jumping in on the game, as indicated by the steep rise in the total number of countries hosting malicious links in Figure 25. Between the entire year of 2008 and the first half of 2009, the number of countries increased by 80 percent.

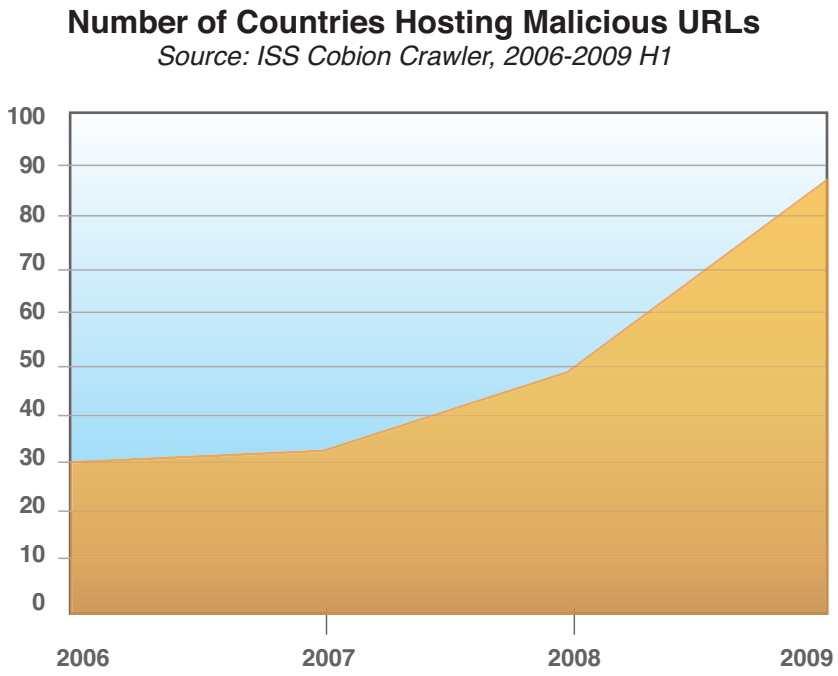


Figure 25: *Number of Countries Hosting Malicious URLs, 2006-2009 H1*

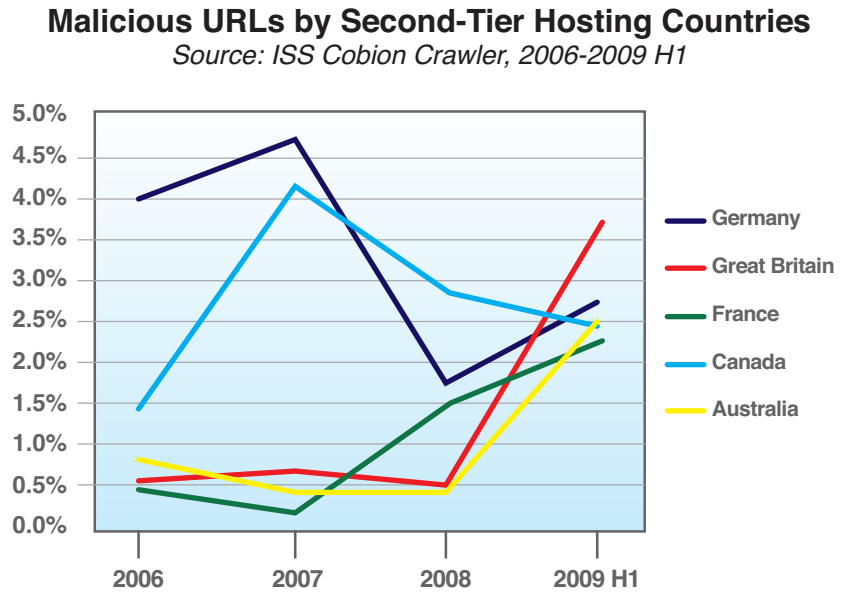


Figure 26: *Second-Tier Countries that Host Two Percent or More of All Malicious URLs, 2006-2009 H1*

Good Web Sites with Bad Links

As described in Web Application Attacks on page 17 and in Common Domains in URL Spam on page 65, attackers are focusing more and more on using the good name of trusted Web sites to lessen the guard of end users and attempt to obfuscate their attempts from protection technologies. The use of malicious Web content is no different. The following analysis provides a glimpse into the types of Web sites that most frequently contain links to known, malicious Web sites. Some of the top categories might not be surprising. One might expect pornography to top the list. However, the second tier candidates fall into the more "trusted" category. Search engines, blogs, bulletin boards, personal Web sites, online magazines and news sites fall into this second-tier category. Most of these Web sites allow users to upload content or design their own Web site, such as personal content on a university's site or comments about a "purchase" on a shopping Web site. In other words, it is unlikely that these types of Web sites are intentionally hosting malicious links. The distribution is probably more representative of the types of Web sites that attackers like to frequent in hopes of finding a loop-hole (like a vulnerability or an area that allows user-supplied content) in which they can incorporate these malicious links in hopes of compromising an unsuspecting victim.

The following chart shows the most common types of Web sites that host at least one link that points back to a known malicious Web site:

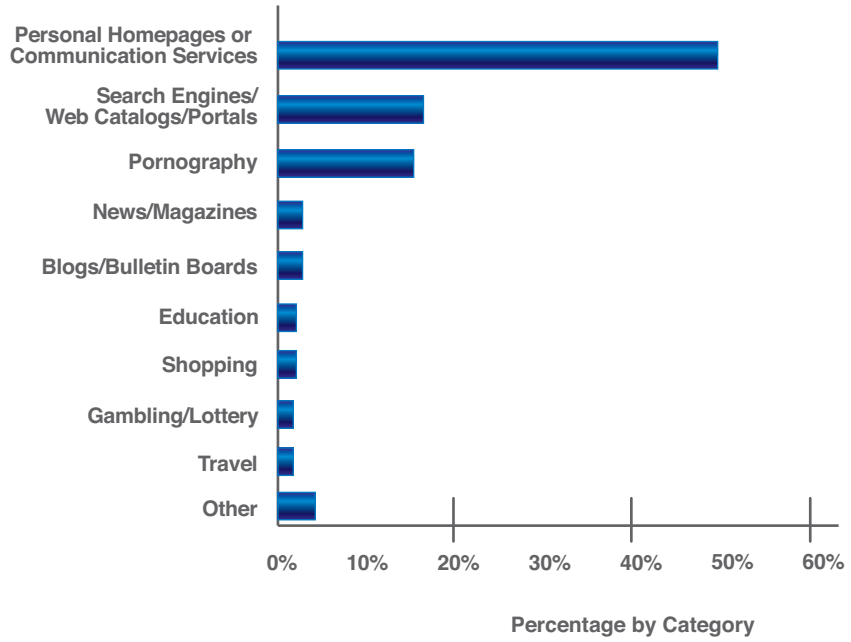


Figure 27: *Top Web Sites Containing at Least One Malicious Link, 2009 H1*

Another way to look at this problem is to examine Web sites that appear to be hosting an extraordinary number of links back to malicious Web sites. When you do an analysis of those that host ten or more links back, another story emerges… one that might imply that the owners of some of these Web sites may be partaking in the financial advantage that these compromises would provide. Out of the categories of Web sites that host ten or more of these links, pornography accounts for nearly 28 percent and gambling accounts for more than 14 percent. One might suspect that these kinds of Web sites are knowingly using these links for profit. Some of these Web sites do appear as if these links were placed systematically throughout the site.

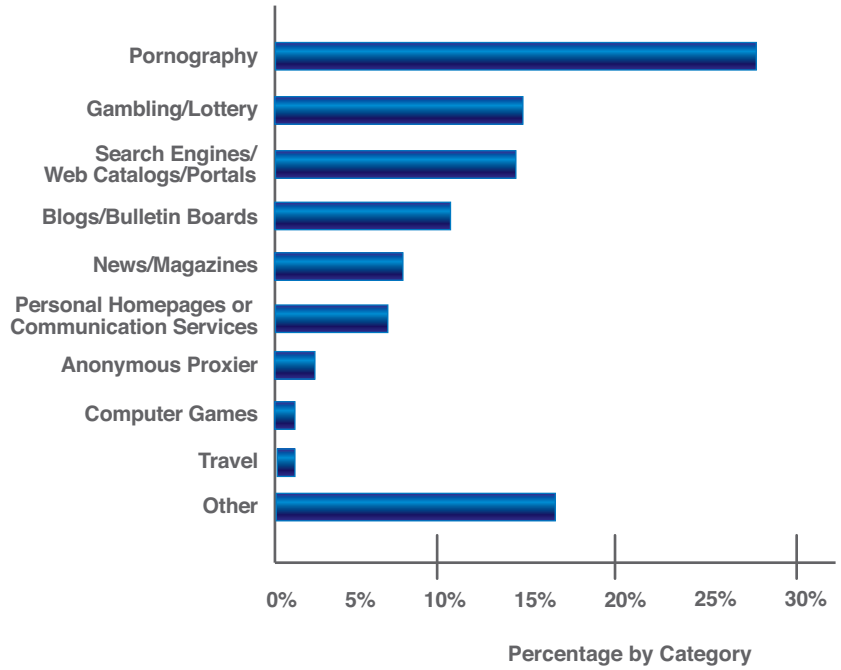


Figure 28: *Top Web Sites Containing Ten or More Malicious Links, 2009 H1*

Malware

Malware Category Trends

Primary Malware Categories

For the first half of 2009, 55 percent of new malware [1] in our collection are Trojans while Backdoors ranked second at 21 percent. Comparing last year's annual report against the first half of 2009, the distribution of Trojans increased by nine percentage points, up from 46 percent.

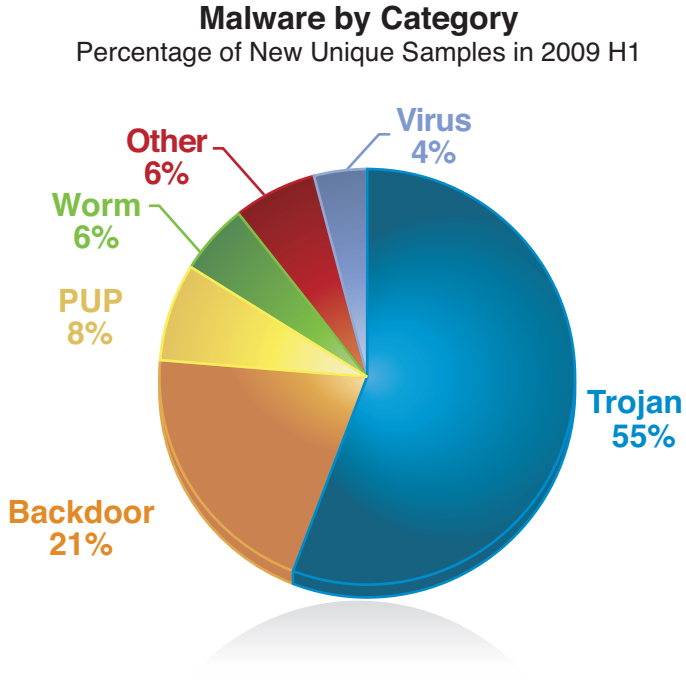


Figure 29: *Malware by Category, 2009 H1*

1 *Samples in our collection are counted as distinct by unique MD5. The percentages in each category represent the total number of unique samples falling into that category which measures the variation of malware in each category, but not the global distribution or propagation of any given sample or family.*

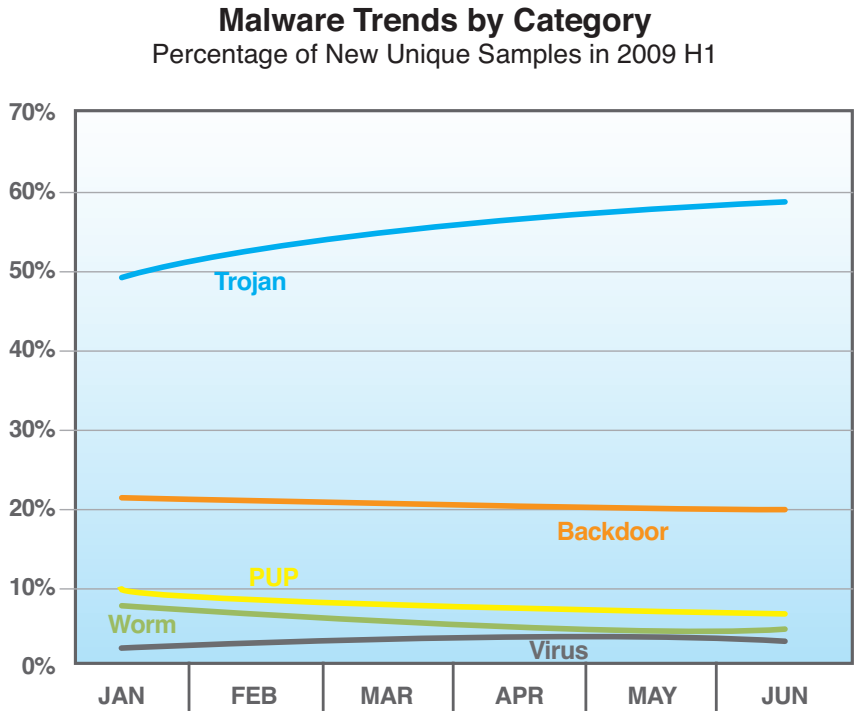


Figure 30: *Malware Trends by Category, Percentage of New Unique Samples in 2009 H1*

Category	Description
Virus	Propagates by infecting a host file
Worm	Self-propagates via e-mail, network shares, removable drives, file sharing or instant messaging applications
Backdoor	Provides functionality for a remote attacker to log on and/or execute arbitrary commands on the affected system
Trojan	Performs a variety of malicious functions such as spying, stealing information, logging key strokes and downloading additional malware
Potentially Unwanted Programs (PUP)	Programs which the user may consent on being installed but may affect the security posture of the system or may be used for malicious purposes. Examples are Adware, Dialers and Hacktools/"hacker tools" (which includes sniffers, port scanners, malware constructor kits, etc.)
Other	Unclassified malicious programs not falling within the other primary categories

Table 11: *Malware Category Descriptions*

One major factor that would explain the high number of Backdoors and especially Trojans in terms of new malware collected is that a large number of new malware today is generated by publicly-available toolkits. The majority of these toolkits are geared toward generating Backdoors (such as Hupigon and Bifrose) to control infected machines and Trojan-Infostealers (such as ZBot, LDPinch and various keyloggers) to spy and steal information from infected machines. This trend is expected to continue since these toolkits are very easy to use, and from a malicious user's perspective, he/she just needs to get the "job" done without much technical investment on their part – these cybercriminals want to just fill out some text boxes, check some check boxes, and then have a pre-configured Backdoor or Infostealer ready in a few seconds. And thus, we would expect to see very similar classification statistics at the end of 2009.

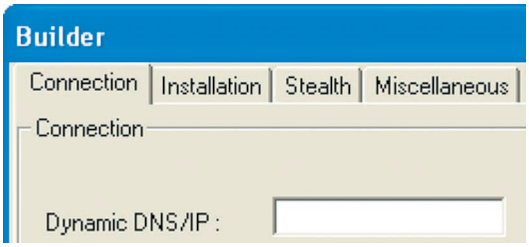


Figure 31: *Toolkits for generating Backdoors and Trojan-Infostealers are popular amongst cybercriminals due to their general availability and ease of use*

In terms of self-propagating malicious programs, the Autorun family tops the ranking in terms of new samples in the Worm category. Incidentally, Conficker (also known as Kido and Downadup) also appears high in the Worm category (due to the number of its minor variants which resulted from re-packing/obfuscation of the major variants) and that one of Conficker's infection vector is also the Windows Autorun feature. Due to the proliferation of malware propagating through the Autorun feature, Microsoft made a move in February of this year and released [2] a security advisory encouraging customers to install an update that fixes an issue which prevents the Autorun feature from being properly disabled.

Figure 32: *In the height of the prevalence of malwares propagating via the Autorun feature (which includes Conficker); Microsoft released a security advisory encouraging users to install an update to the Autorun feature of Windows.*

2 http://www.microsoft.com/technet/security/advisory/967940.mspx

Additionally, Waledac (also known as Iksmas) also ranked high in our Worm category. Waledac's main purpose is to build a spam botnet, and it uses spam and social engineering to spread. Waledac was first seen in the wild last December. Since then, it has used various social engineering ploys ranging from Christmas e-card to terrorist bombing spam themes to tricking users into downloading a copy of the malware from the Web links that are embedded in the spam e-mails.

Trojan Category Breakdown

As the Trojan category continues to have the largest share in terms of new samples in our collection, we provide the breakdown of the Trojan category below.

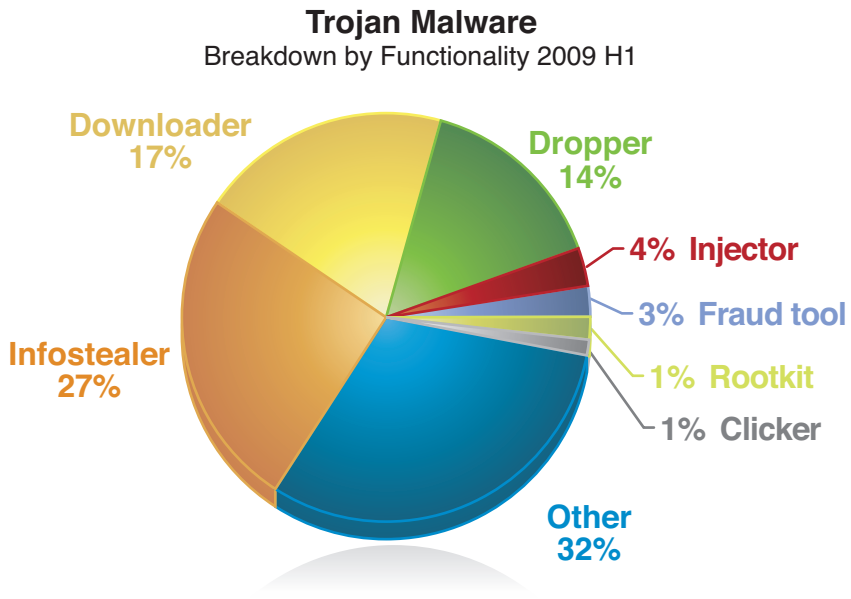


Figure 33: *Trojan Category Breakdown, 2009 H1*

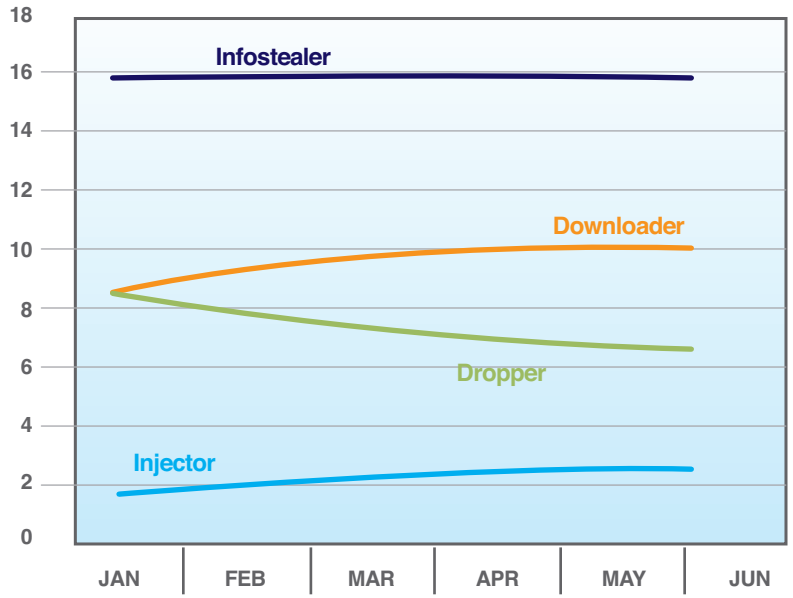


Figure 34: *Trojan Trends, 2009 H1*

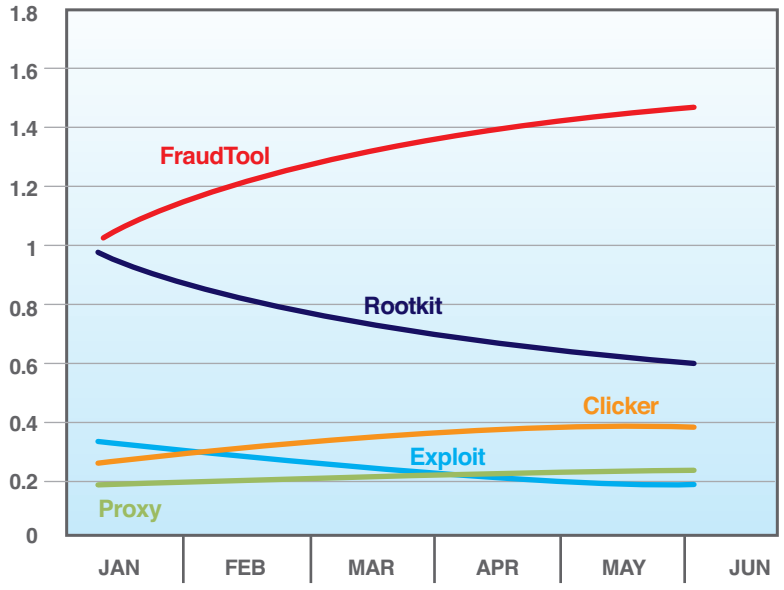


Figure 35: *Trojan Trends, Granular Detail for Other Category, 2009 H1*

Category	Description
Infostealer	Spies and/or steals information. This category includes password stealers, keystroke loggers and spyware.
Downloader	Downloads one or more malware components from a remote site and then installs them on the affected system.
Dropper	Drops and installs one or more embedded malware components onto an affected system.
Injector	Injects an embedded malware component into other another process. One purpose is for the embedded (and usually obfuscated) malware to evade antivirus detection. Another purpose is for the embedded malware to evade host-based firewalls by injecting it into a trusted process such as a browser or a system process.
FraudTool	Malware used to commit fraud. An example is malware that displays fake error or infection messages, and then entices the user to purchase fake tools or security software.
Clicker	Generates website traffic, the purpose of which is to generate revenue or other malicious purposes.
Rootkit	Components used by other malware in order to have the capability to hide themselves from the user and security software.
Exploit	Documents or media files containing exploit code.
Proxy	Allows a remote attacker to relay connection through the affected system in order to hide its real origin.
Other	Trojans that do not fall within the other subcategories.

Table 12: *Trojan Category Descriptions*

Similar to the primary categories breakdown, the Trojan category breakdown was very similar to the result of our annual report for 2008. A major fraction of new Trojans samples are categorized as Infostealers (27 percent), followed by Downloaders (17 percent) and Droppers (14 percent). We also added a new Trojan subcategory called Injector. Using another mechanism to install a malware payload into the affected system, Injectors encompass malicious programs which inject their embedded malware payload into other processes as an attempt of the embedded payload to evade antivirus detection and/or evade host-based firewalls.

For the first half of 2009, we continued to see a high number of Infostealer Trojans targeting online game users (Trojan-Infostealer.Onlinegames, Trojan-Infostealer.Magania, Trojan-Infostealer.Tibia and Trojan-Infostealer.Wow). These types of Infostealers numbered around 5 percent of our total collection. Likewise, Microsoft reported [3] that the two top threats removed by its Malicious Software Removal Tool (MSRT) in February were online game password stealers, thereby further illustrating their prevalence. Infostealers and Downloaders targeting online banking users (Trojan-Infostealer.Banker, Trojan-Downloader.Banload and Trojan.Infostealer.Zbot) also continue to be in the top ranking Trojans in terms of new samples.

In addition, Trojan-Downloader.FraudLoad, a downloader family whose main purpose is to download and install rouge security software is also among the top-ranking Trojan families for which we collected the most new samples.

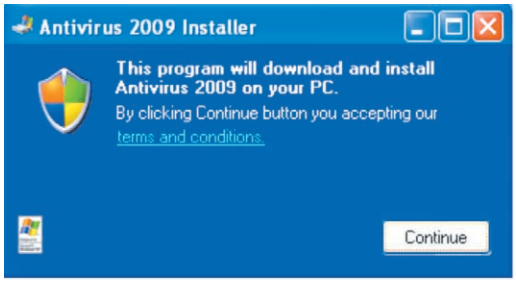


Figure 36: *An example dialog box shown by a variant of Trojan-Downloader.FraudLoad. Even if the user does not click the "Continue" button, a rogue security software will still be downloaded and installed*

On the other hand, Trojan-FraudTool, our categorization for rogue security software (scareware), is still a small percentage of our collection (around 1.5 percent for the first half of 2009); however, our data shows a small but continued increase in this category.

3 http://blogs.technet.com/mmpc/archive/2009/02/19/msrt-observations-online-game-password-stealers.aspx

Looking at the top ranking Trojan families in our collection, we can see where cybercriminals are likely to be profiting. The reason is that new samples can indicate activity on the part of the cybercriminals, and activity (such as generating new samples in order to evade detection, change configuration or to upgrade functionality) may indicate incentive and continued profit in a chosen area. And thus, in the case of the first half of 2009, we can conclude that there is continued profit in selling stolen online game credentials and virtual assets, selling/using stolen online banking credentials, and scamming users into buying fake security software.

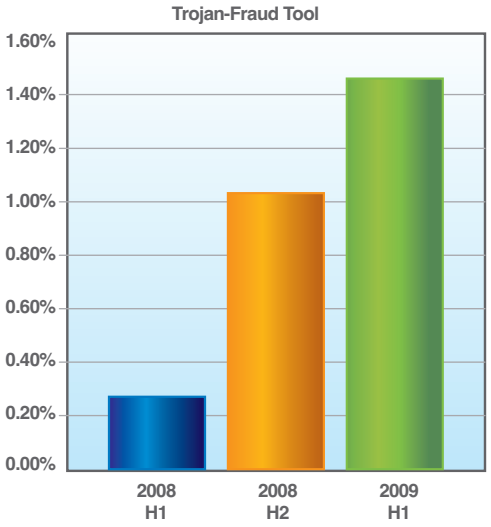


Figure 37: *Trojan-FraudTool Trend (2008 H1–2009 H1, percentage of total samples collected)*

Top Phone Home Locations

In 2008, we had begun the process of recording malware activity information from the malware samples that we are collecting. Malware activity is collected by running the samples through the Virus Prevention System (VPS) engine – IBM ISS' behavior-based antivirus engine. Through its virtualization capability, we are able to generate reports containing information such as the changes made in the file system, registry, and service database. We are also able to record information relating to the malware's network activity such as downloading activity and connecting to/sending data to specific IP addresses or host names.

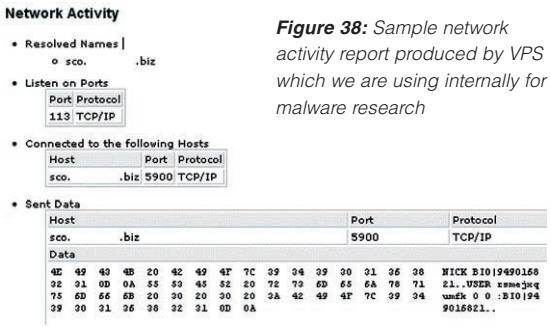


Figure 38: *Sample network activity report produced by VPS which we are using internally for malware research*

One of the interesting statistics we had pulled from the virtualization results is a list of phone home IP addresses or host names which can include the addresses where the malware either downloaded files from, or sent data to or received data from for its Command and Control (C&C) communication.

The figure below shows the geographical distribution of malware phone home locations from the samples we collected for the first half of 2009:

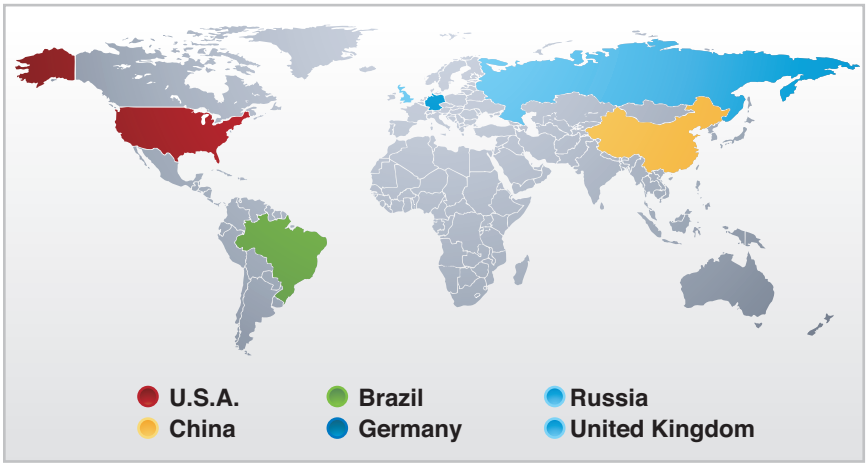


Figure 39: *Phone Home Locations Geographical Distribution (Unique IPs), 2009 H1*

The table below shows the Top 5 countries where malware phone home locations are located:

Rank	Country	Percent
1	USA	35%
2	China	14%
3	Brazil	8%
4	Germany	4%
5	Russia and United Kingdom	3%

Table 13: *Top Phone Home Locations, 2009 H1*

Based on our data, the USA hosts a large percentage (35 percent) of phone home machines, followed by China (14 percent) and then Brazil (8 percent).

Additionally, TCP port 80 is the most common phone port used, and it is mostly used for downloading via HTTP and transferring information or sending infection notification messages via HTTP GETs and POSTs. However, there is also some malware that use custom protocols but are also using TCP port 80 to pass the data.

There are a couple of reasons why malware authors usually select HTTP as a communication channel. One reason is that it is fairly easy to set up and use (by using URLMON and WinInet APIs). The other reason is to avoid suspicion by using a very common protocol. In relation to the latter, it is interesting to see that a lot of the download URLs for executables files (or sometimes a data file such as a replacement hosts file) have an extension of an image file, mostly .jpg. Simply put, the malware is trying to make the update look like an image request over the Web, probably one of the most frequent types of Web traffic on the Internet.

```
http://www.    .net/users/ca    cce/svchosts.jpg
http://www.    .net/users/ca    cce/spoolv.jpg
http://www.    .net/users/ca    cce/msnmsrg.jpg
```

Figure 40: *Examples of suspicious download URL where the extension is that of an image file (.jpg) but the basename is similar to those of an executable file*

Another interesting fact that we collected is that some malware uses FTP as a phone home channel. The malware author embeds the FTP credentials in the malware body and if another malicious user obtains a copy of the malware, he/she can easily hijack the data that the malware may have stored in the FTP server.

Conficker: Story and a Lesson Learned

It is to no surprise that Conficker is the main story for the first half of 2009. From its silent beginnings in December to the height of its fame in April, Conficker had baffled security researchers, caused panic among computer users, and had shown us a glimpse of the mindset and the sophistication of cybercriminals. This section presents the story of Conficker and the lesson we had learned from it.

Conficker Started Small

The Conficker problem we faced in the first half of 2009 was actually seeded in the wild last year, in late November [4]. It started as a simple network worm which spreads via the Windows Server Service vulnerability (MS08-067) that was announced by Microsoft just a month before. The cybercriminals did not waste any time to take advantage and infiltrate machines not yet patched against the vulnerability. However, Conficker did not become largely prevalent until the end of December when a new variant (Conficker.B) started to incorporate additional propagation mechanisms including the propagation through network shares and removable drives. The prevalence of Conficker then led to the development of the Conficker Working Group in February followed by an announcement by Microsoft for a reward to those who can help identify the Conficker author/s.

Researchers Baffled

In the meantime, after the release of Conficker.A and Conficker.B, security researchers were still baffled as to what the purpose of Conficker would be, and why the malware author/s did not include any commands in Conficker's code to perform any other activities other than to spread as fast as possible and to download an executable file from several Web sites. No one really knew what Conficker's next move would be until these Web sites were activated by the malware author/s. The mystery was further complicated because the Web site addresses that Conficker will connect to is based on 250 generated domain names that change daily, which also made it difficult to block access to these Web sites.

```
vrycbs.org
kxqkll.info
lvepxprfndg.info
zmxnz.biz
lkfxcolccno.biz
woheqsphqt.biz
vniidcsog.net
esqtmc.net
gspdekw.net
fsoikxwy.info
yaoqat.net
qbkibhshenu.biz
kjbryf.org
aisvgblqnc.com
xkgnuv.info
vpgwmwqh.info
keoyi.biz
uckpxb.com
pbpfoiuf.net
kjquwg.biz
```

Figure 41: *Example domain names generated by Conficker. They appear to be random but Conficker's author/s knows what domain names will be generated at specific dates.*

4 http://www.confickerworkinggroup.org/wiki/pmwiki.php/ANY/Timeline

This domain name generation feature of Conficker was proven to be effective. In March, Conficker.C was reported [5] to have been successfully downloaded by Conficker.B-infected machines from a number of the download Web sites. Unfortunately, the Conficker Working Group was not able to control the registration of some of the domain names that Conficker could generate.

P2P Botnet Capability Unveiled

Meanwhile, while all this was unraveling in the early parts of March, we had begun noticing an increase in UDP traffic in one of our darknets. The UDP packets seemed to be directed to random UDP ports and contained what seemed to be random data with random lengths. Several days after, we received a copy of Conficker.C and began the process of dissecting it. What resulted from the dissection of the Conficker.C was a surprise – the malware author/s removed Conficker's propagation routines and somewhere hidden in the code was a capability that the malware author/s had taken special care not to be easily understood by researchers – it was Conficker's peer-to-peer (P2P) botnet communication routine. Another surprise came when we found out that it is the same code that generated the random UDP traffic that we had seen earlier in March.

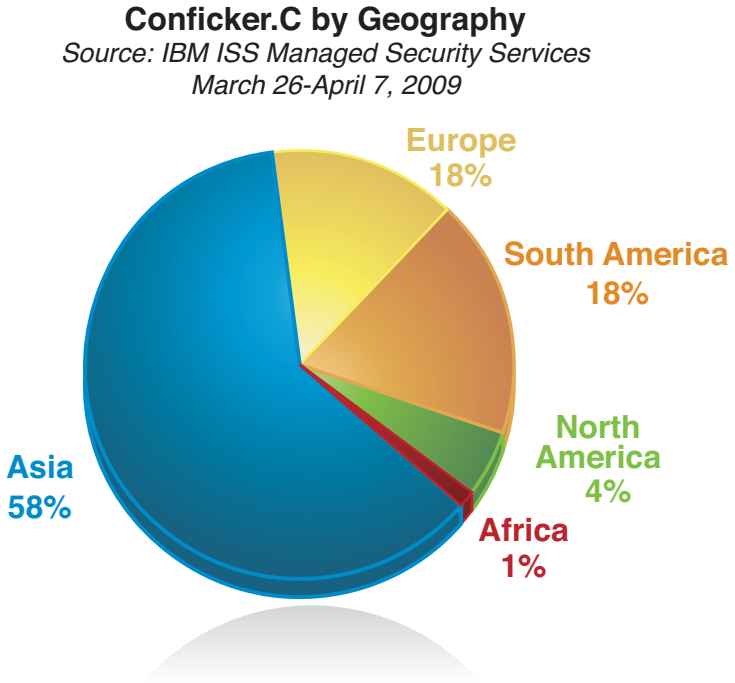


Figure 42: *Regional distribution statistics of Conficker.C a few days after we released the Conficker.C P2P detection signature*

5 http://www.confickerworkinggroup.org/wiki/pmwiki.php/ANY/FAQ

Rank	Country	Percent
1.	China	16.6%
2.	Brazil	10.8%
3.	Russian Federation	10.2%
4.	Republic of Korea	4.6%
5.	Vietnam	4.5%
6.	India	4.1%
7.	Ukraine	3.6%
8.	Indonesia	2.9%
9.	Italy	2.9%
10.	Taiwan	2.8%

Source: *IBM ISS Managed Security Services, March 26-April 7, 2009*

Table 14: *Country distribution ranking of Conficker.C a few days after we released the Conficker.C P2P detection signature*

After several days of analysis, we were able to decode Conficker's P2P communication protocol and create an IDS/IPS signature to passively identify Conficker.C-infected machines. Once our Conficker.C P2P detection signature was deployed, it became clear to us the extent of the infection and the distribution of Conficker.C around the world. It was fascinating that suddenly, we were able see where Conficker.C-infected machines were, whereas before they seemed to be just random chatters in the Internet that no one could understand.

At that point, it was clear that Conficker was now gearing up to build a P2P botnet infrastructure in which the malware author could distribute any executable code he wants. The other important point to understand is that this P2P botnet has no central command thereby making it difficult to shut down and track down the controller of the botnet. This architectural decision was again a carefully planned action by the cybercriminals; similar to the domain name generation technique that they had implemented, the P2P botnet infrastructure they are building has no single point of failure.

The "April Fools Computer Worm"

But of course, the story did not end in March. Conficker.C has one more trick up its sleeve; it was again set to start downloading an executable file from several Web sites starting on April 1. However, this time, Conficker.C would be downloading the file from Web sites in which the addresses were based on domain names that are taken from a pool of 50,000 generated domain names that changed daily. The upcoming trigger date plus a combination of media attention caused panic as the dubbed "April Fools Computer Worm" would start receiving instructions from the cybercriminals on April 1.

However, nothing happened as April 1 passed, reminding us of the "Michelangelo Madness" [6] that happened March 1992 in which the Michelangelo boot virus was predicted to be wiping hundreds of thousands, then million of systems on its trigger date of March 6, 1992. But when the date came, the reports were proven to be inflated.

Monetizing the Botnet

In Conficker's case, the malware author/s did not do anything on April 1; instead, they had another plan. It is was several days after they started acting on their plan when several security firms noticed that a new version of Conficker (Conficker.D/E) and rogue security software were being distributed through the Conficker P2P network and not in the download Web sites where these payloads were anticipated to be distributed.

It then became clear what the main purpose of the Conficker is – Conficker was created by the cybercriminals as a platform for mass distributing any executable content they want – it can be an updated version of Conficker, and more importantly monetize this distribution platform by distributing other types of malware, and in the case in April, it was rogue security software.

---

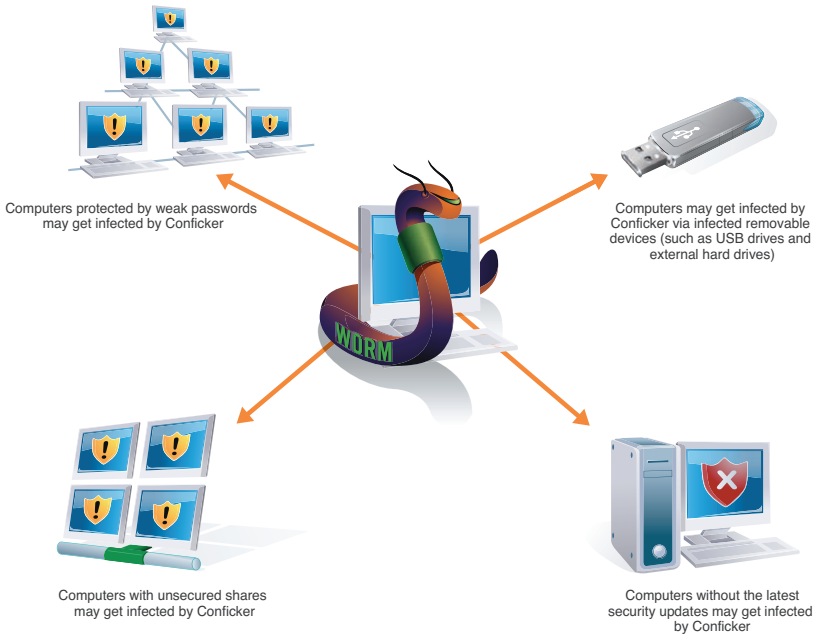
6 http://www.research.ibm.com/antivirus/SciPapers/White/VB95/vb95.distrib-node7.html

The story still continues because a large number of machines are still infected today and, at any given time, the cybercriminals can distribute any new executable payload on these infected machines. Only time will tell what the cybercriminal's next move is, and we can just hope that law enforcement agencies will identify the people responsible for Conficker before they can make their next move.

Lesson Learned

Looking back, we think of Conficker as a test of an institution's security posture and that the prevalence [7] of Conficker and the reported high-profile victims ranging from universities to government agencies tell us that a lot of institutions failed the test.

Blended threats such as Conficker will try to infiltrate systems using a number of possible means and if there is a weak link in the security chain, the whole chain will be broken. As an example, a workstation with all the latest software updates, protected by a strong password, with updated antivirus software, and with firewall/IPS enabled would still get infected if the user inserted a USB drive which, unbeknownst to the user, is infected by a new, undetected Conficker variant – game over.



7 http://www.confickerworkinggroup.org/wiki/pmwiki.php/ANY/InfectionTracking

Conficker is just one example of the malware threats facing us today and it did not even use all the weapons in a malware author's arsenal in order infiltrate our systems and protect its foothold. Having learned from Conficker and the ones before Conficker, we can predict that the malware authors will create even more sophisticated malware.

What if in the week or two, the ultimate blended attack arrives: a malware capable of infiltrating systems through the exploitation of human trust (via social networks, fake Web sites, spam e-mails, instant-messaging and file sharing networks), exploitation of vulnerable system components (such as operating system and browser-based vulnerabilities), and finally, through the exploitation of weak policies (via removable devices and systems protected by a weak password)… can you say that you are ready for it? Unfortunately, we think the answer from most organizations would be "no."

Spam

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

Currently, the spam filter database contains more than 40 million relevant spam signatures (every spam is broken into several logical parts [sentences, paragraphs, etc.], and a unique 128-bit signature is computed for each part) and millions of spam URLs. Each day there are one million new, updated or deleted signatures for the spam filter database.

The topics of this section are:

- *New trends around spam types*
- *Most popular domains used in spam*
- *Most popular Top Level Domains (TLDs) used in spam and why the top domains are so popular*
- *Lifespan of spam URLs*
- *Spam's country of origin [8] trends, including spam Web pages (URLs)*
- *Changes in the average byte size of spam*
- *Most popular subject lines of spam*
- *Recovery from the McColo Takedown*

---

8 The statistics in this report for spam, phishing, and URLs use the IP-to-Country Database provided by WebHosting.Info (http://www.webhosting.info), available from http://ip-to-country.webhosting.info. The geographical distribution was determined by requesting the IP addresses of the hosts (in the case of the content distribution) or of the sending mail server (in the case of spam and phishing) to the IP-to-Country Database.

Spam Volume

The spam volume has not evolved and expanded as in years past. Instead of a steady increase, spam flattened out near the middle of last year with a significant drop in November due to the McColo takedown (see Recovery from the McColo Takedown on page 78 for more information about McColo). In the beginning of 2009, spam volume stagnated for a couple of months, and then started to increase in May, finally reaching (and surpassing) the spam level seen just before the McColo shutdown last year.

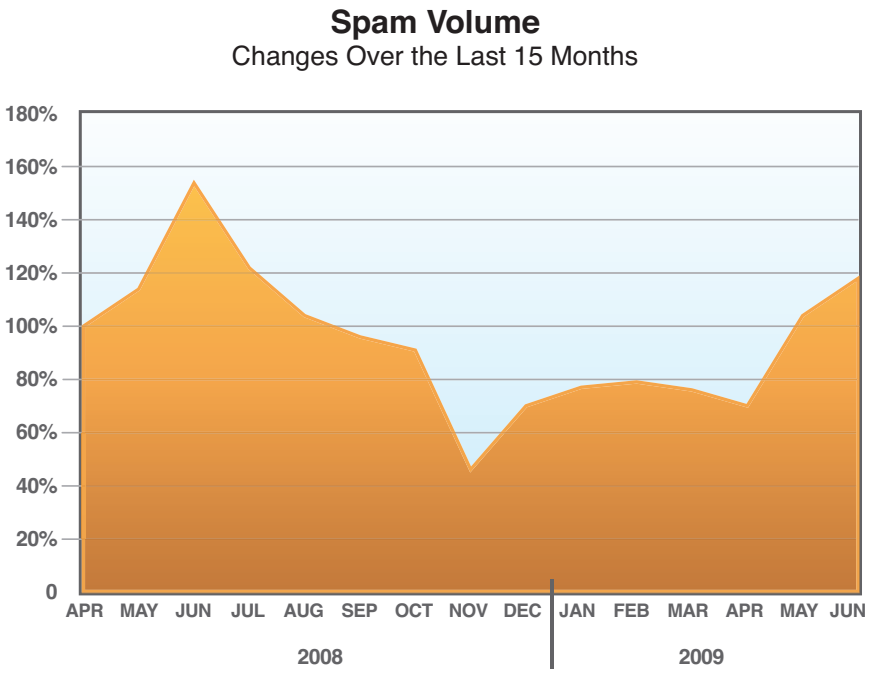


Figure 43: *Changes in Spam Volume since April, 2008*

Types of Spam

In 2008, spammers focused on using the most unsuspicious type of e-mail: HTML-based spam without attachments. The chart below shows a significant increase concerning this type. However, in the second quarter of 2009, two trends emerged. Spammers changed their strategy and started creating more single, plain-text spam (without other e-mail parts or attachments), and we also witnessed the rebirth of image-based spam:

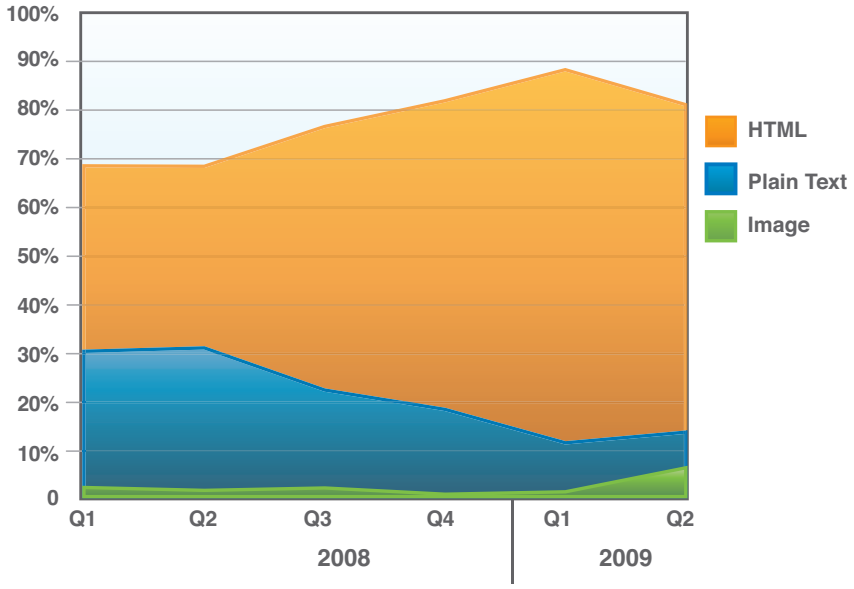


Figure 44: *Types of Spam, 2008-2009 H1*

The Rebirth of Image-Based Spam

Image-based spam boomed in 2006 and 2007, but practically disappeared in 2008 except for a short stint in October of that year.

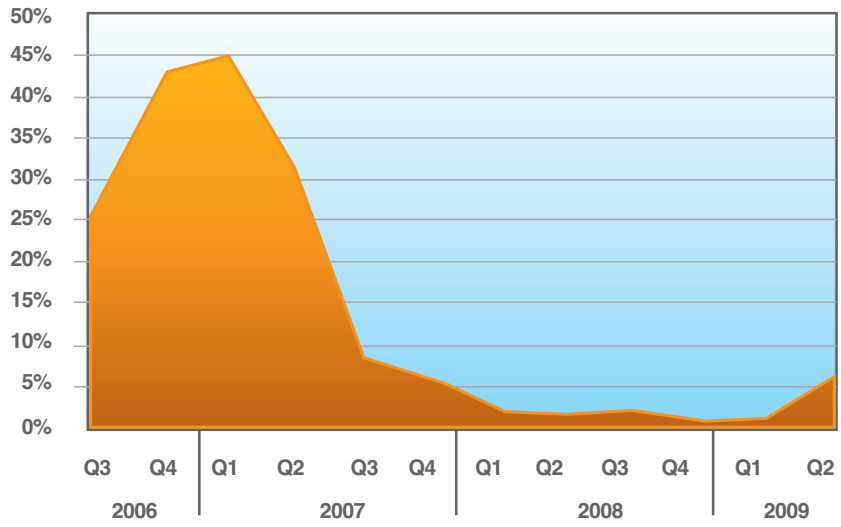


Figure 45: *Image-Based Spam 2006 Q3 until 2009 Q2*

Shortly before the McColo shutdown, image-based spam made a brief appearance, and then shortly stopped after the shutdown in November of 2008 took its toll.

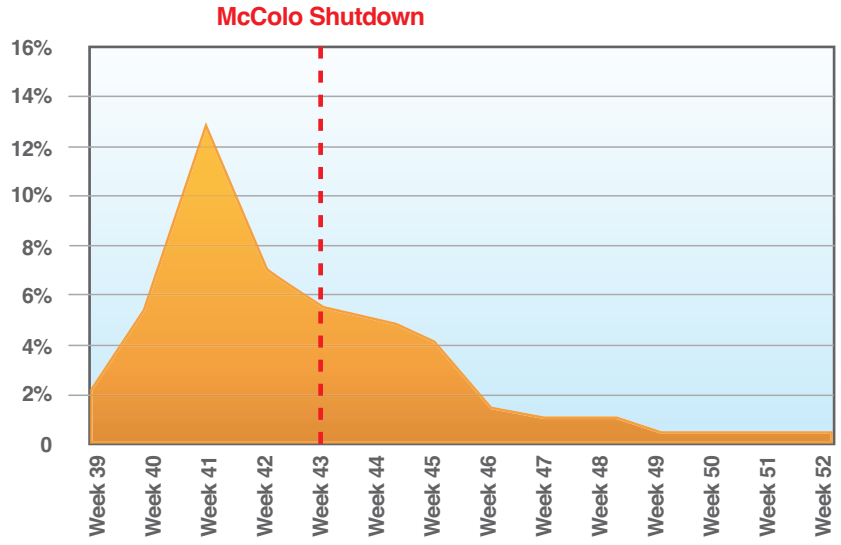


Figure 46: *Image-Based Spam 2008 Q4*

Image spam was down another four months, but, then, in March of this year, spammers started three new runs of image-based spam, and the third one was still running its course at the end of June:

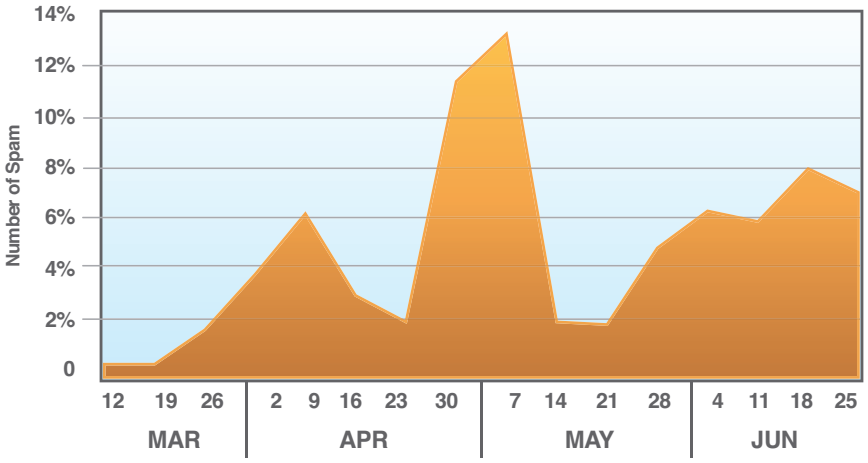


Figure 47: *Image-Based Spam March to June, 2009*

Here are a few details about the trial runs sent in March and April:

- *Most of them are of pharmaceutical nature, advertising drugs, pills, etc.*
- *Only a few of them use random pixels, and many of them even have identical binaries.*
- *Many of these spam messages contain random text below the image.*
- *Most of them do not contain any Web links that the user can click.*
- *Most of them ask the user to visit a .com Web site with a domain name consisting of six digits like 123456.com, and the user has to manually type that URL into the browser.*

Technically, there were no new techniques in this spam. Thus, most anti-spam filters should block them, for example, by using fingerprints (like IBM Proventia Network Mail Security System and IBM Lotus Protector for Mail Security do).

From the WHOIS information of the domains shown on the images, all of them have similar WHOIS registration information. The domains are registered at registrars that are infamous for URL spam, like:

- *35 TECHNOLOGY CO., LTD*
- *CENTROHOST CLOSED JOINT STOCK COMPANY*
- *XIAMEN ENAME NETWORK TECHNOLOGY D/B/A ENAME.CN ENAME.COM*
- *XIN NET TECHNOLOGY CORPORATION*

Regarding the content of the spam, there was only one major difference in comparison to the image spam of 2007. Two years ago, most spam focused on stock trading. With the financial crisis happening, stock spam probably isn't a lucrative option for spammers. The focus on drugs is possibly an attempt at preying on people that want to "feel better" during desperate times.

So, why would the spammers return to an old technique, especially when getting a successful bite requires a user to actually type the URL into the browser themselves? Perhaps they were trying to mask their URLs through these images. In their trial run near the end of March, did they see that some anti-spam systems were losing their edge when it came to image spam? Are they simply running out of new ideas and rehashing old techniques?

It will be interesting to see what comes next... maybe we will see another resurgence of PDF spam (considering the focus PDFs have received from an exploitation standpoint, it seems likely), MP3 spam, or even spam with hidden, random text (white text on white background).

Have we somehow hit a plateau of spam techniques? Who knows? We can tell you that from the monitoring perspective, it all feels a bit strange. It's like sitting down to watch the storyline progress in your favorite TV show only to find that the directors have inexplicably substituted an 80's-style montage in its place.

Common Domains in URL Spam

The vast majority of spam, 60 percent of it, is still classified as URL spam, spam messages that include URLs that a person clicks to view the spam contents:

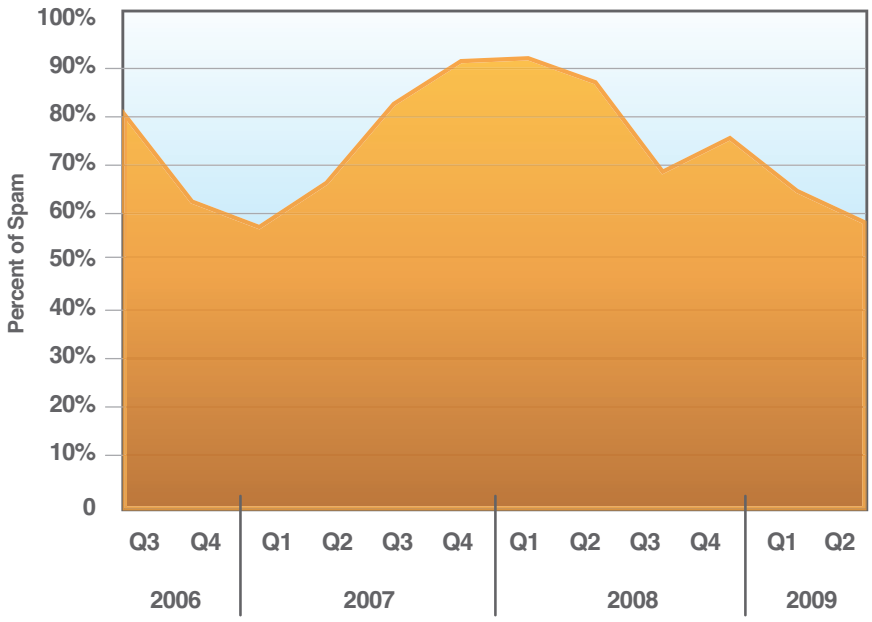


Figure 48: *URL Spam, 2006 Q3 to 2009 Q2*

Hence, it is worthwhile to take a closer look at the most frequently used domain names in URL spam. The following tables show the top 10 domains per month throughout 2008, with a few key domains highlighted.

Rank	January 2008	February 2008	March 2008	April 2008	May 2008	June 2008
1.	googlepages.com	blogspot.com	blogspot.com	crazeben.com	doubleclick.net	dogpile.com
2.	sarahkverok.com	81.222.138.69	powref.com	manninst.com	livefilestore.com	kewww.com.cn
3.	magnarx.com	goldsmallman.com	nuelig.com	hyuaien.com	maddris.com	ynnsuue.com
4.	nesoeteaok.com	fastmansilver.com	gelsedde.com	pobueitah.com	nubteku.com	wpoellk.com
5.	lifefreeart.com	dotoneauto.com	mewlegos.com	congratym.com	moieiaus.com	movecontinent.com
6.	sgmykrtrewt.com	dedeiooss.com	findmilk.com	timeminute.com	coridez.net	moptesoft.com
7.	qualiveok.com	geocities.com	marketthen.com	camethank.com	zimpleq.com	varygas.com
8.	nightboylost.com	hotripefruit.com	seatbar.com	wroteleast.com	misllie.com	earexcept.com
9.	northmanestimate.com	topstopcool.com	believeagree.com	writecotton.com	pogieamdo.com	fullrow.com
10.	geocities.com	fastpetsilver.com	somelisten.com	saveany.com	poskeij.com	colonytop.com

Table 15: *Most Common Domains in URL Spam, 2008 H1*

Rank	July 2008	August 2008	September 2008	October 2008	November 2008	December 2008
1.	livefilestore.com	cnn.net	livefilestore.com	livefilestore.com	live.com	gucci.com
2.	smellshort.com	cnn.com	imageshack.us	live.com	tubdyqwenqe.com	notdune.com
3.	elementdepend.com	msn.com	beroyal.info	el1te-russ1an-g1rls.com	eurocasinokd.com	hereidea.com
4.	opera.com	msnbc.com	forformisskasino.com	myrusfriend.net	stop-fl0p.net	live.com
5.	grayany.com	imageshack.us	totalwrite.com	yellowpages.com	bbc.co.uk	heatdark.com
6.	creasehappiness.com	reoisk.com	cazinoyoumeyou.com	livechatfreex.com	hop-m0p.com	namenot.com
7.	msn.com	google.com	casinonewtrip.com	googlegroups.com	t1p-top.com	idolreplicas.com
8.	boceph.com	soieuu.com	csinomonster.com	cazinosostermor.com	eurocasinokg.com	davavkos.com
9.	alizedup.com	royalfirsteuro.info	beroyal.mobi	777-models-777.com	n1cewomen7.com	vutovlaf.com
10.	augsid.com	royalfirsteuro.mobi	beroyal.org	cazinomonste.com	sexymodels123.net	conemain.com

Table 16: *Most Common Domains in URL Spam, 2008 H2*

Although the majority of URL spam is hosted on domains that were registered for spam purposes, the amount of URL spam using well-known and trusted domain names has continued to increase. In the first half of 2008, these well-known domains made our monthly top ten list only 8 times. In the second half of 2008, this count more than doubled with 19 spots filled with well-known names. In 2009, this trend continues with 31 spots filled.

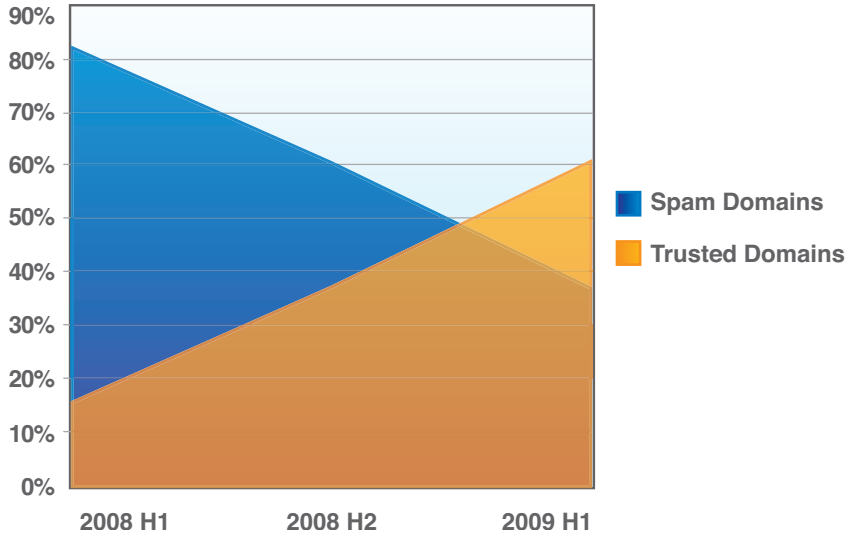


Figure 49: *Top Ten Domains Used in Spam, Spam Domains Versus Trusted Domains, 2008-2009 H1*

The following table highlights the well-known domains falling in the top ten list for the first half of this year. In March and April, 8 and 9 of the top 10 used domains in spam were well-known domains.

January 2009	February 2009	March 2009	April 2009	May 2009	June 2009
chat.ru	sexyhardy.com	rodale.com	interia.pl	yahoo.com	yahoo.com
thuspattern.com	aspirationask.com	menshealth.com	akamaitech.net	menshealth.com	googlegroups.com
powerinstrument.com	shoprespect.com	webmd.com	menshealth.com	icontact.com	webmd.com
cbsnews.com	msn.com	mkt41.net	ask.com	webmd.com	icontact.com
hereidea.com	yulesearching.com	interia.pl	webmd.com	earlytorise.com	mansellgroup.net
notdune.com	wordobservant.com	icontact.com	rodale.com	doctorspreferred.com	ranmooon.com
methoddegree.com	assistingoriginal.com	akamaitech.net	go.com	mansellgroup.net	signgras.com
chithigh.com	tarecahol.cn	msn.com	yahoo.com	healthcentral.com	rannew.com
chitlink.com	integrityprove.com	about.com	yimg.com	menshealth.fr	blueheav.com
boughtprosperity.com	approvaltruthful.com	rodalenews.com	behaviorright.com	trendsmag.com	rangreat.com

Table 17: *Most Common Domains in URL Spam, 2009 H1*

Some of the well-known Websites are:

- *about.com (online source for original information and advice, owned by The New York Times Company)*
- *akamaitech.net (Web site of Akamai Technologies)*
- *ask.com (internet search engine)*
- *cnn.com (official Web site of the Cable News Network, owned by Time Warner)*
- *go.com (web portal, operated by the Walt Disney Internet Group)*
- *googlegroups.com (free service from Google where groups of people have discussions about common interests)*
- *healthcentral.com (official Web site of The HealthCentral Network, medical information portal)*
- *icontact.com (e-mail marketing offering company)*
- *interia.pl (large Polish Web portal)*
- *mansellgroup.net (official Web Site of Mansell group, a marketing services company)*
- *menshealth.com (official Web Site of Men's Health Magazine, published by Rodale Inc.)*
- *msn.com (a joint venture between NBC Universal and Microsoft for online news)*
- *rodale.com (official Web Site of Rodale Inc., publishes health and wellness magazines, books, and digital properties)*
- *webmd.com (official Web Site of WebMD Health Corporation, an American provider of health information services)*
- *yahoo.com (Major Internet search engine)*

Not only do these legitimate Web sites provide a recognizable (and trustworthy) Web link to the end user, but spam messages using them may also successfully evade some anti-spam technology because they only use legitimate links in their spam e-mails.

Common Top Level Domains in URL Spam

The Top Level Domain .com dominates the domain table in the previous section. However, the analysis of Top Level Domains reveals another story of what sparks the interest of spammers. The following tables show the five most frequently used Top Level Domains used in spam by month:

Rank	January 2009	February 2009	March 2009	April 2009	May 2009	June 2009
1.	com	com	com	com	com	com
2.	cn (China)	cn (China)	cn (China)	cn (China)	cn (China)	cn (China)
3.	org	org	org	pl (Poland)	org	org
4.	ru (Russia)	ru (Russia)	net	net	net	net
5.	net	net	pl (Poland)	org	ru (Russia)	ru (Russia)

Table 18: *Most Common Top Level Domains in Spam, 2009 H1*

This table shows the Top Level Domains used within spam independent from the availability of the corresponding Web sites. When considering only the Top Level Domains of those URLs that really host spam content then we have:

Rank	January 2009	February 2009	March 2009	April 2009	May 2009	June 2009
1.	com	com	cn (China)	com	cn (China)	cn (China)
2.	cn (China)	cn (China)	com	cn (China)	com	com
3.	ru (Russia)	ru (Russia)	ru (Russia)	at (Austria)	ru (Russia)	net
4.	net	net	net	in (India)	net	ru (Russia)
5.	es (Spain)	es (Spain)	at (Austria)	org	fr (France)	org

Table 19: Most Common Top Level Domains with real Spam content, 2009 H1

The – maybe surprising – result is that the most spam content is not hosted on .com Domains but on .cn Domains, at least in March, May, and June, 2009. As in previous years, the only purpose of including .com Domains (which were typically randomly-generated and not even accessible or functioning URLs anyway) in spam is to make look them more legitimate. Using .com URLs in spam is the most unsuspicious type of URL because 55 percent of all domains used on the Internet are .com domains (source: IBM ISS data center, see Web Content Trends on page 32 for more details).

Country Code Top Level Domains (like .cn, .ru, .es) are not used randomly. Nearly 100 percent of those URLs do really host spam content (or redirect to spam content automatically) if they are used in a spam message, which is different for the Generic Top Level Domains (like .com, .net, and .org). The following chart shows TLDs that most frequently use random Domains (without hosting spam content).

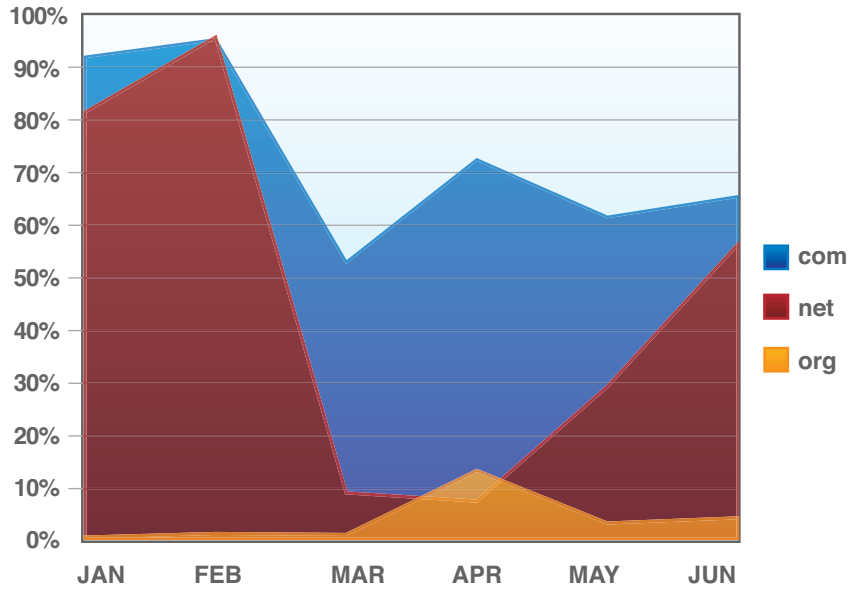


Figure 50: *Percentage of URLs per TLD that Host Real Spam Content – 2009 H1*

As the chart shows, .org URLs found in spam e-mails are typically these randomly-generated, fake URLs. Others, like .net URLs tend to fluctuate from month to month. This trend in .com Domains started in the past six months. It will be interesting to see whether hosting spam content on .com URLs continues to decline throughout the rest of the year.

Lifespan of Spam URLs

Over the past few years, the URLs that these spam messages point to have had a shorter and shorter lifespan. The quicker they are put up and taken down, the more likely they will avoid detection. Three years ago, more than half of the URLs used in spam were up for longer than a month. In the last 12 months, more than 95 percent of these URLs were up for a week or less. In the last quarter, nearly 99 percent of them are up a week or less as shown in Figure 51.

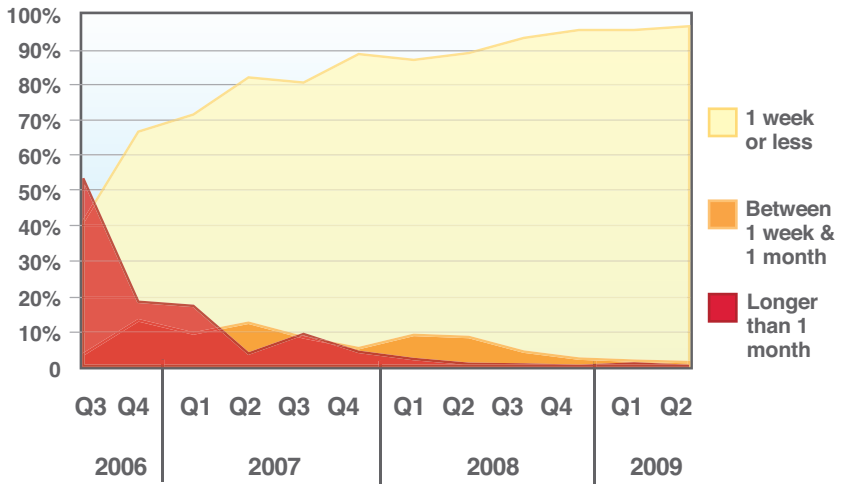


Figure 51: *Lifespan of Spam URLs 2006 Q3 through 2009 Q2*

Since the lifespans are continually getting shorter, the following chart shows a breakdown of the lifespan of 2009 spam URLs in terms of days.

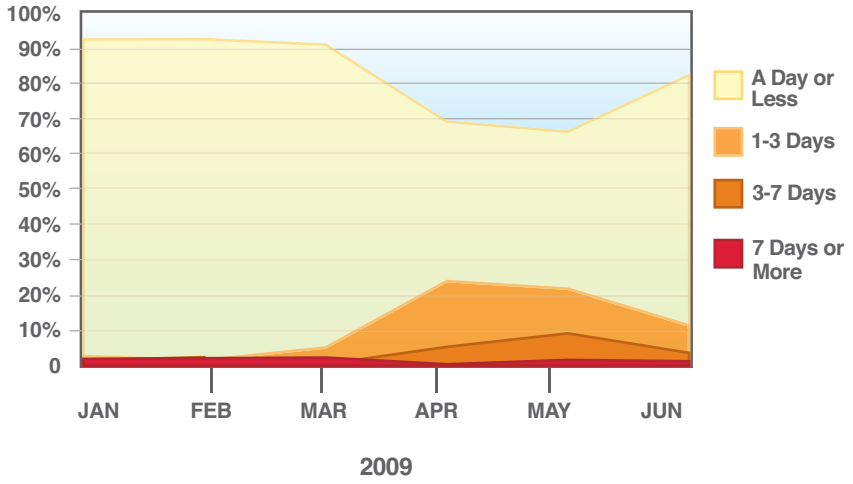


Figure 52: *Lifespan of Spam URLs in Days, 2009 H1*

Figure 52 provides some interesting data points on URL spam. An astonishing percent (generally, 70 to 90 percent) of URLs are live for 24 hours or less. Although the percentage appears to drop in April and May, there is an eerie correlation here—the resurgence of image-based spam. It appears that at least some of the short-lived URL spammers, during April and May, switched from using 24 hour or less URL spam to using image-based spam.

Spam—Country of Origin

The following map shows the origination point for spam globally in 2009 H1. Brazil, the U.S., and India account for about 30 percent of worldwide spam.

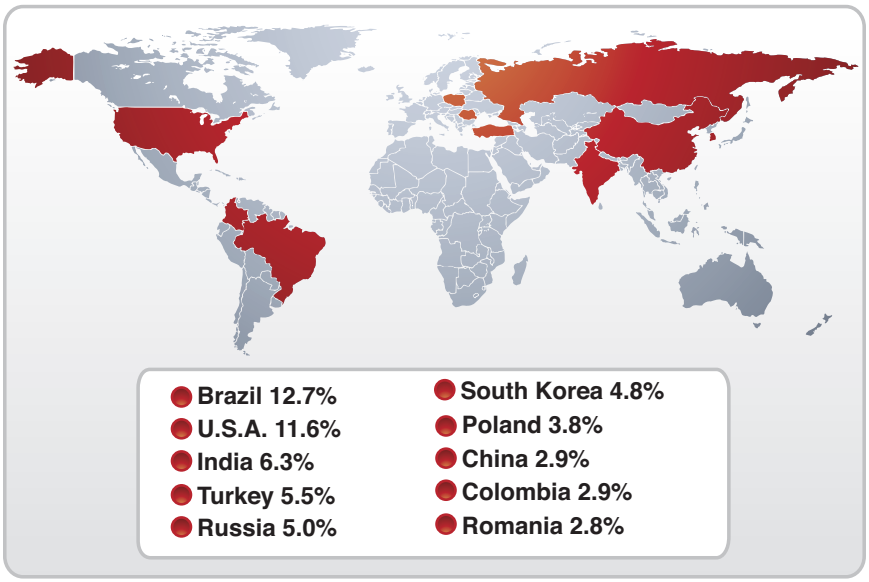


Figure 53: *Geographical Distribution of Spam Senders, 2009 H1*

Spam—Country of Origin Trends

There are two newcomers in the top three countries from which spam originates: Brazil and India. After the McColo shutdown, India was one of the countries that bounced back the fastest, surpassing their original quantity of spam before the end of 2008. So, it appears as if their "success" has continued bringing them to the top three.

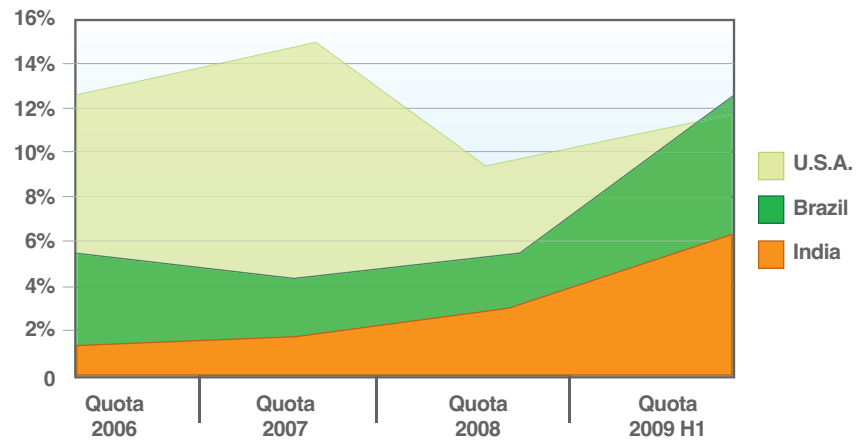


Figure 54: *Spam Origins Over Time: Brazil, India and the US, 2006-2009 H1*

Growth in BRIC Countries

Brazil and India, as the third and the forth BRIC country, have shown rapid growth in the spam and phishing industries. The other two BRIC countries, Russia and China, have not been complacent in this regard. Russia is the top country for the origin of phishing e-mails, and China is the top hosting country for spam URLs. For BRICS, spam and phishing are two industries that are experiencing rapid growth alongside many other industries in these countries.

Spam URLs—Country of Origin

The following map shows where the spam URLs are hosted.

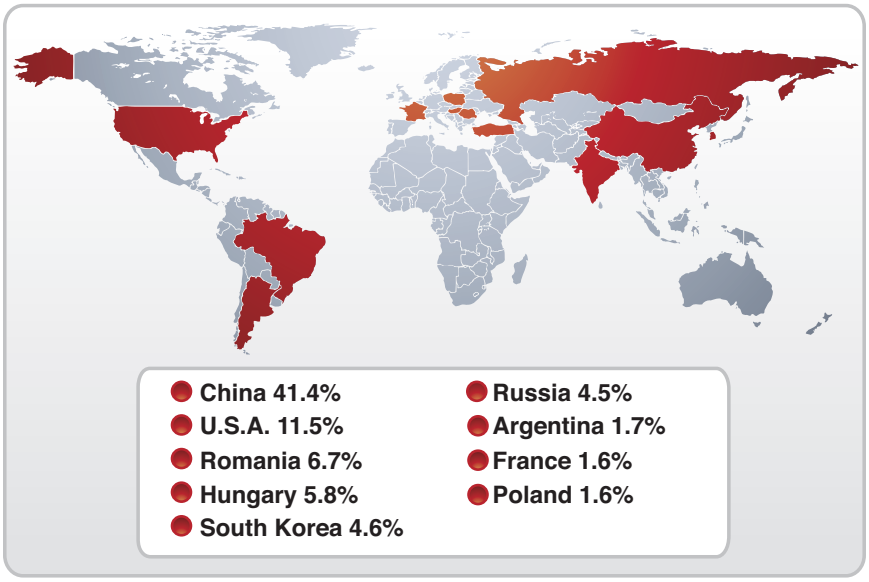


Figure 55: *Geographical Distribution of Spam URLs, 2009 H1*

Spam—Average Byte Size

The most significant change in the average byte size of spam happened at the end of 2007 and corresponded with the decline of image-based spam. In 2008, byte size began to rise ever so slightly up until the McColo takedown later in the year. With the resurgence of image-based spam, the last months have seen a resurgence in the average size of spam, too. The average size exceeded 5 kilobytes for the first time in one and a half years.

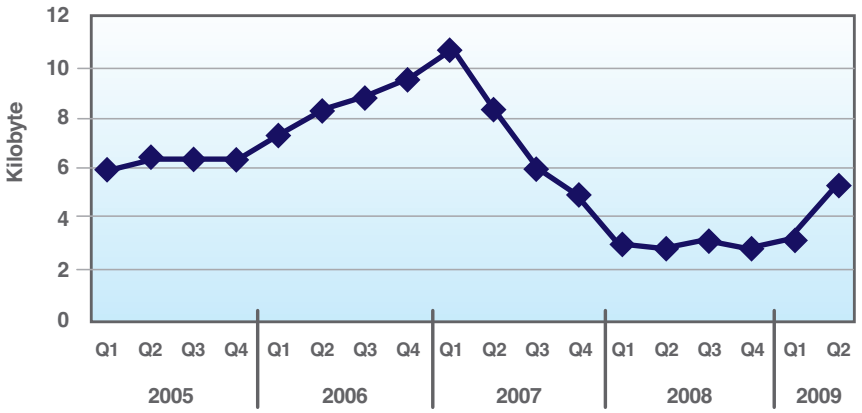


Figure 56: *Average Byte Size of Spam since 2005*

Spam—Most Popular Subject Lines

Whilst spam subject lines became more and more granular from 2007 to 2008 this trend is reversing slightly in the first half of 2009. The top ten subject lines in the first half of 2009 make up about 5.6 percent of all spam subject lines, up from 3 percent in 2008, but still down from the 20 percent figure recorded in 2007.

In the first half of 2009, the percentages of the top 10 subject lines increased significantly. As shopping on the Internet becomes more and more popular, spammers use subjects about an order's status to attract the user's interest. Furthermore, offers of replica watches and e-mail that appears to be sent by online sales companies like Amazon are very often used to attract the user's attention.

The following table shows the most popular spam subject lines in 2009 H1:

Subject Lines	%
You've received an answer to your question	0.76%
Swiss Branded Watches	0.71%
Customer Receipt/Purchase Confirmation	0.69%
Email Handling Opinion Needed	0.68%
Hi	0.67%
Replica Watches	0.58%
Great Finds	0.41%
Check out hot deals	0.39%
Exquisite Replica	0.37%
Sales Receipt Amazon	0.35%

Table 20: *Most Popular Spam Subject Lines 2009 H1*

Recovery from the McColo Takedown

After the takedown of the California-based Web hoster McColo in November of last year, the spam volume dropped to around 25 percent of previous levels. The sudden and extreme volume and country distribution changes observed after the shutdown demonstrated that McColo was the base operator of spam bots all around the world.

Changes in International Distribution of Spam

The United States has, for years, maintained a top spot in the spam origin list. Six days before the takedown, it was in the number one spot. Six days after the takedown, spam production coming out of the US was reduced to a mere 14 percent of its original capacity. So, it was not a terrible surprise when the US finally lost its top spot on the list on this sixth day after the takedown.

Top Spammers Before and After the McColo Takedown

Just Before		Just After		End of 2008		2009 Q1		2009 Q2	
USA	14.2%	China	12.7%	Brazil	11.7%	Brazil	12.4%	Brazil	13.2%
Russia	11.0%	Russia	11.4%	USA	8.1%	USA	11.2%	USA	12.3%
Turkey	7.4%	USA	8.0%	China	6.6%	India	6.6%	India	5.8%
Spain	5.9%	South Korea	6.2%	Turkey	5.7%	Turkey	5.3%	Turkey	5.8%
Brazil	4.8%	Brazil	5.8%	Russia	5.7%	Russia	5.2%	South Korea	5.5%

Table 21: *Top Spammers Before and After the McColo Takedown*

So, has the US recovered from the McColo takedown? Almost. In the first two quarters of 2009, Brazil was the top spam sender, and the US held the second position. Both Brazil and the US increased their overall percentage and the distance from the third "competitor," India.

Phishing

This section covers the following topics:

- *Phishing as a percentage of spam*
- *Phishing country of origin trends, including phishing Web pages (URLs)*
- *Most popular subject lines and targets of phishing*
- *Phishing targets (by industry and by geography)*

Phishing Volume

Throughout 2008, phishing volume was, on average, 0.5 percent of the overall spam volume. In the first half of 2009, phishing attacks have decreased dramatically to only 0.1 percent of the spam volume. We know that the criminal networks behind phishing use methods for identity theft other than sending out a simple e-mail that looks like a legitimate e-mail coming from a bank. The decline in phishing and increases in other areas (such as banking Trojans) indicate that attackers may be moving their resources to other methods to obtain the gains that phishing once achieved.

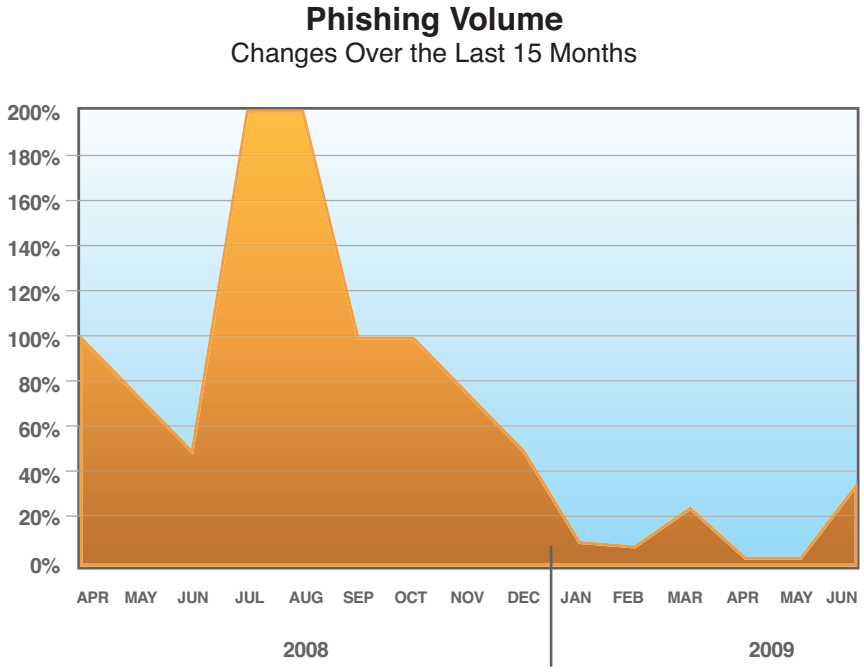


Figure 57: *Phishing Volume, Apr 2008-Jun 2009*

Phishing—Country of Origin

Along with the dramatic changes in phishing volume come other dramatic changes, like the country of origin. Spain and Italy took slots one and two in 2008, but Spain has completely dropped from the top ten for the first half of 2009. The top sender now is Russia, who was not even in the top ten last year. Other changes include the addition of Turkey, Ukraine, and India and also the disappearance of Israel, France, and Germany, who were smaller players in 2008.

The following map highlights the major countries of origin for phishing e-mails in 2009 H1.

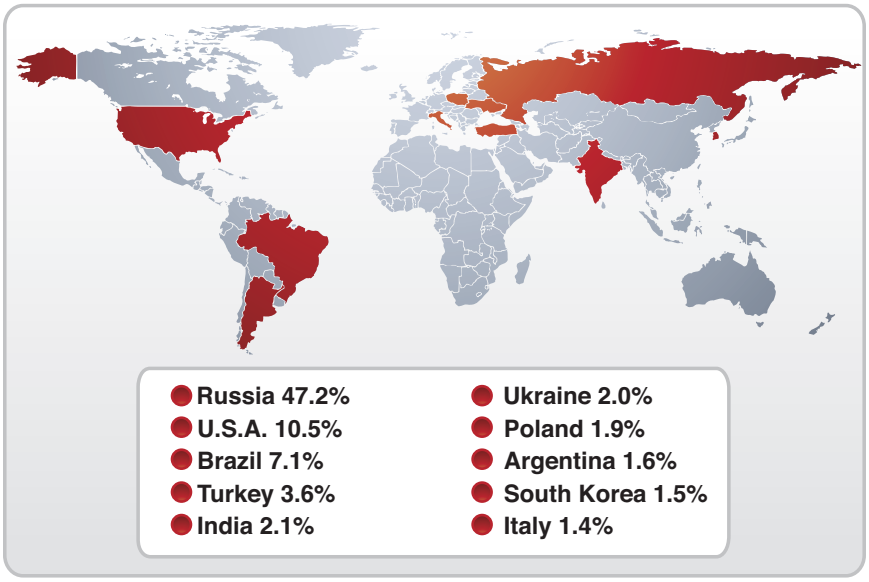


Figure 58: *Geographical Distribution of Phishing Senders, 2009 H1*

Phishing URLs—Country of Origin

The following map shows where the phishing URLs are hosted. The top nine players have not changed in comparison to 2008, although their place in the top nine has changed slightly in some cases. In the tenth position, Poland has replaced Thailand.

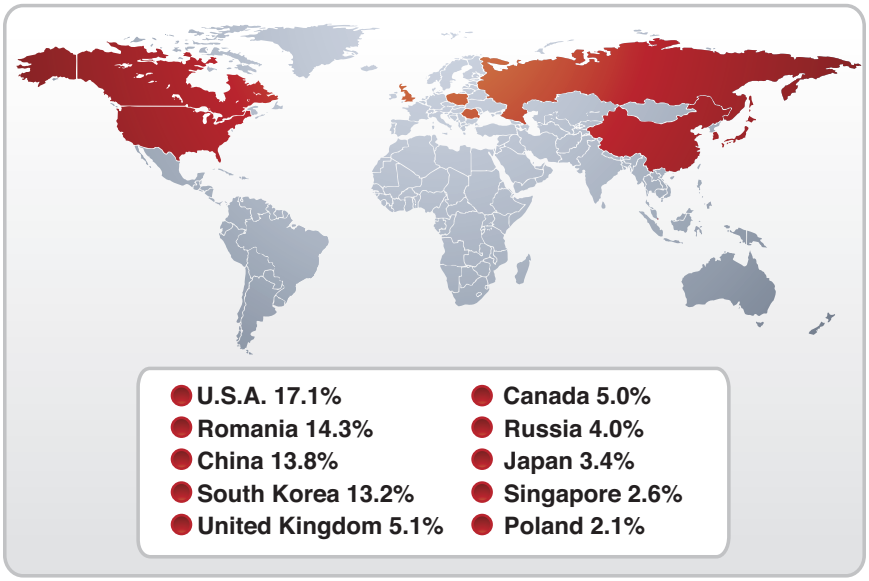


Figure 59: *Geographical Distribution of Phishing URLs, 2009 H1*

Phishing—Most Popular Subject Lines

One of the biggest changes in 2008 was that popular subject lines were not so popular anymore. In 2007, the most popular subject lines represented more than 40 percent of all phishing e-mails. In 2008, the most popular subject lines made up only 6.23 percent of all phishing subject lines. Thus, phishers became more granular in their targets in 2008, essentially with a greater variance of subject lines than in 2007.

In the first half of 2009, the trend was reversed completely when it comes to phishing subject lines: The top 10 most popular subject lines represent more than 38 percent of all phishing e-mails. Similar to 2008, the most popular subject lines are dominated by PayPal.

It is interesting that the subject holding the number one slot is from a PayPal phishing e-mail directed at French speakers. However, the subject is incorrectly spelled because of missing accents (the correct version would be: Attention! Votre compte PayPal a été limité!). It appears that Phishers have limited French language skills, not only because of the misspelling, but also because there is only one subject variation for French PayPal phishing e-mails. For the English versions, there are three PayPal variations in the Top 10.

Another possibility is that their phishing kits have not gone through a mandatory I18n (Internationalization) process like the rest of us in the software industry, and so their kits are simply limited to 7-bit characters, excluding characters like é, à etc.

The following table shows the most popular phishing subject lines in the first half of this year:

Subject Lines	%
Attention! Votre compte PayPal a ete limite!	24.05%
Important Information Regarding Your Limited Account.	7.02%
PayPal® Account Review Department	2.06%
Account Security Measures	1.35%
Citibank Alert: Additional Security Requirements	1.33%
Important Information Regarding Your Account.	0.89%
Online Account Security Measures	0.53%
PayPalŽ Account Review Department	0.5%
Paypal Account Update	0.44%
Security alert	0.27%

Table 22: *Most Popular Phishing Subject Lines 2009 H1*

Phishing Targets

Phishing—Targets by Industry

In 2008, financial institutions were unquestionably the dominant target of phishing e-mails. In the first half of 2009, financial institutions are still the number one target. Along with the decline in phishing and the change in phishing origins, the actual targets of phishing have changed significantly. Financial institutions now only represent 66.3 percent of the targets, allowing Online Payment institutions to consume 31.4 percent of the share. This change in percentage is not necessarily indicative of more phishing directed towards Online Payment organizations, but more accurately represents the decline in North American and European financial targets when it comes to phishing.

The other 2.3 percent of phishing targets is comprised of other industries such as online auction Websites, communication services, and online stores:

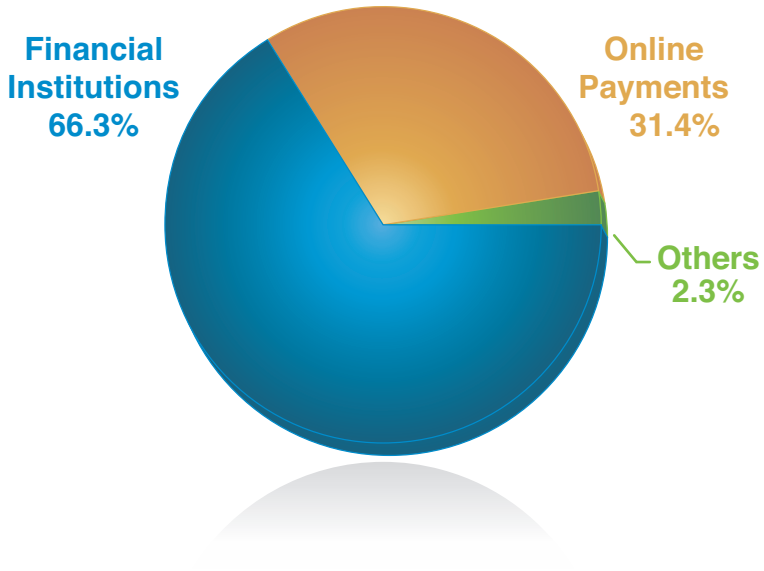


Figure 60: *Phishing Targets by Industry, 2009 H1*

This distribution of targets is a major change in comparison to 2008, as the following chart demonstrates.

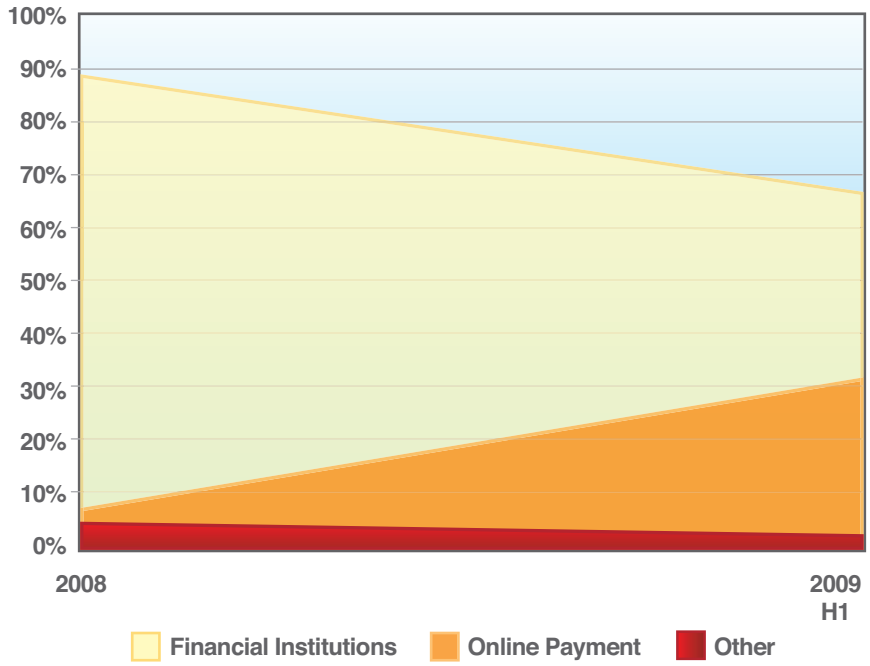


Figure 61: *Phishing Targets by Industry, 2008-2009 H1*

Now that nearly 50 percent of all phishing e-mails are sent from Russia, the fact that more online payment phishing is coming from Russia is not surprising. The "traditional phishing e-mails" targeting online banking services are in decline, perhaps because of the financial crisis or because of improved security measures for verifying that you are indeed logging into the real online bank. Since online banking fraud is by no means decreasing, phishers must be finding other ways to compromise users instead of the traditional phish in which they send out e-mails that look like they came from a bank (as mentioned above). However, online payment phishing still seems to be lucrative.

Phishing—Financial Targets by Geography

Over 99 percent of all financial phishing targets are still in North America or Europe. The overall numbers for the first half of this year are similar compared with last year. In 2008, 58.4 percent were directed at North American institutions and 40.8 percent at European institutions. In the first half of this year, the majority of targets are still in North America (nearly 65 percent) and Europe is a second runner-up with 35 percent:

Figure 62: *Financial Phishing by Geographical Location, 2009 H1*

However, after taking a closer look using shorter time frames, dramatic changes become more apparent. The following chart shows the shift in geographical location of the two major financial institution phishing targets (North America and Europe) that happened over the course of this year so far.

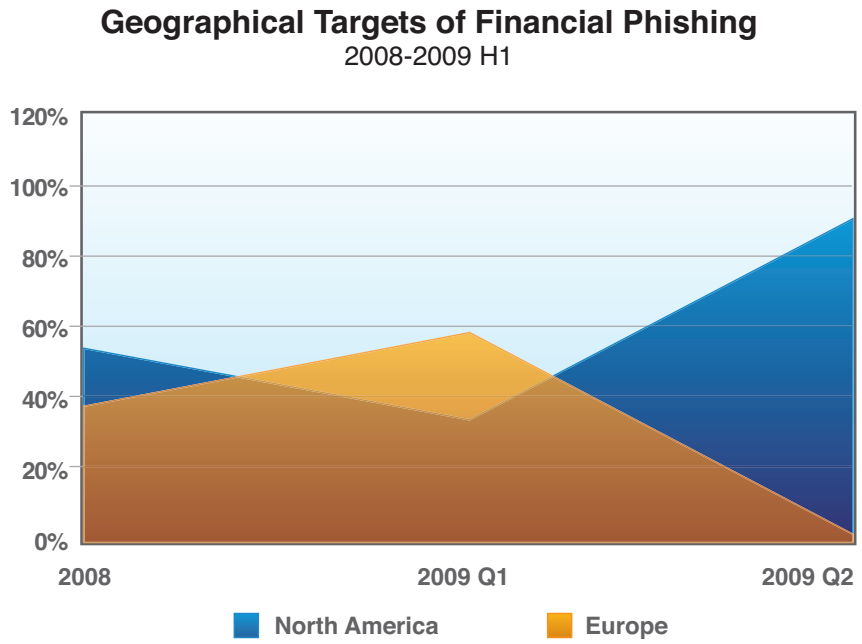


Figure 63: *Financial Phishing by Geographical Location, 2008 – 2009 Q2*

In the first quarter of 2009, European financial phishing targets exceeded North American targets. In the second quarter, phishing targeted at European banks has become nearly obsolete. Perhaps phishers that are switching to other bank fraud techniques (like malware) tend to specialize in European markets.

Perhaps they feel that North America is rebounding from the financial crisis faster than Europe, and so they are refocusing their sights on North American banks. If this were true, then the different consequences of the financial crisis for North America and Europe would seem to be reflected in the financial phishing e-mails.

IBM Global Services

Route 100

Somers, NY 10589

U.S.A.

Produced in the United States of America.

August 2009