Outline

• Spam
• Viruses
• Worms
• Spyware

• [Worms in depth]

SPAM

• Problem
  – Zero marginal cost of sending an email

• Solutions
  – Machine learning client to detect spam
    – Brightmail
      • Dummy accounts
      • Correlate SPAM messages
      • Supply fingerprint to enterprise customers
  – Client refuses messages from unknown senders, until
    • They respond to a Turing test query
    • They execute a computationally expensive applet
    • Micropayment

Link Spam

• Keyword / Meta tag stuffing
  – Linguistic spoofing

• Multiple titles

• Tiny fonts

• Invisible text
  – <body bgcolor="FFFFF">Your text here</body>
  – Problem: takes up space. Size=1? Bottom?

• Doorway / jump pages
  – Fast meta refresh

• Cloaking – Code swapping

• Pagerank spoofing (Link newtworks)

Robots

• Threat: automatic creation of accounts
  – Paypal
  – Storage associated: Hotmail, Yahoo communities…
  – Adbots in chat rooms
  – Online polls

• Solutions
  – Turing tests
    • Distorted speech recognition
    • Overlayed distorted text recognition
    • CAPTCHA
      – Automated public Turing test to tell computers and humans apart
      – http://www.captcha.net/

Gimpy: Type 3 words

Mori & Malik (UCB) program solving ez-gimpy with accuracy 83%
### Viruses

- **Defn**
  - Requires human action to spread
  - Infects most files on local computer
  - Doesn’t automatically spread across network
  - Carries payload (destructive or annoying messages)

- **Common Modus Operandi**
  - Macro attached to office document

- **Solutions**
  - Fingerprint based (to detect viruses)
  - Application checksums (to detect tampering)

### Worms

- **Defn**
  - Automatically spreads to other systems

- **Modus Operandi**
  - Protocol worms
  - Hybrid virus / worms

- **Solutions**

### Spyware: analysis and mitigation

**Steven Gribble**
Department of Computer Science and Engineering
University of Washington

**Kingsofchaos.com**

- A benign web site for an online game
  - earns revenue from ad networks by showing banners
  - but, it relinquishes control of the ad content
Incident: January 2004

• kingsofchaos.com was given this “ad content”

• This “ad” initiated a cascade of redirections through many sites, and ultimately:
  – bombarded the user with pop-up ads
  – hijacked the user’s homepage
  – exploited an IE vulnerability to install spyware

What’s going on?

• The advertiser was really an ad-spammer
  – his goal: force users to see ads from his servers
  – revenue from ad “affiliate programs”
  – paid to show ads for bogus anti-spyware software

• Why install spyware?
  – to show ads whether or not the victim is on the Web
  – to make escape hard
  – his spyware shows his ads
  – the hijacked home page shows his ads
  – some of his ads re-install spyware and re-hijack

Take-away lessons

• Your PC has value to third parties
  – spyware tries to steal this value from you
    • adware: eyeballs and demographic information
    • spyware: sensitive data, PC resources

• Web content should never be trusted
  – even if its direct provider is

• Consumer software and OSs are weak
  – browsers are bug-ridden
  – OSs cannot deal with malicious software

What is Spyware?

• Incredibly difficult to define precisely
  – no clean line between good and bad behavior
  – hard to define “informed consent”

• Spyware is a software parasite that:
  – collects info of value and relays it to a third party
  – hijacks functions or resources of PC
  – installs without consent of user, resists de-installation

• Spyware provides value to others, but not to you

Types of spyware

<table>
<thead>
<tr>
<th>Class</th>
<th># signatures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cookies and web bugs</td>
<td>47</td>
</tr>
<tr>
<td>Browser hijackers</td>
<td>272</td>
</tr>
<tr>
<td>Adware</td>
<td>210</td>
</tr>
<tr>
<td>Keyloggers</td>
<td>75</td>
</tr>
<tr>
<td>Dialers</td>
<td>201</td>
</tr>
<tr>
<td>Backdoors / monitors</td>
<td>279</td>
</tr>
</tbody>
</table>

From the “Spybot S&D” database, Feb. 2005

Infection methods

• Piggybacking on legitimate software
  – provides revenue stream for free software vendors

• Drive-by downloads
  – malicious Web content exploits browser vulnerability
  – software is installed and run silently

• Installed during remote attack
  – some worms now carry spyware payload

• Snowball effect from existing spyware
  – trojan downloaders
Spyware trends

• Most Internet PCs have it
  – June ’03: 80% of Internet-connected PCs are infected
    – [AOL/NCSA online safety study]
• It’s getting more vicious
  – December ’04: 14% of enterprise PCs have backdoor or monitor spyware
    – doubled between October ’04 and December ’04
    – [Webroot reported scan statistics]
• Convergence of threats
  – worms, viruses, spyware, botnets are fusing

Two research studies

• November 2003 study of adware within UW
  – passive network measurement of entire campus
  – measured spread of four adware programs
• Sneak preview of crawler-based study
  – active retrieval of content from the Web
  – how much is “out there,” and who is spreading it

UW adware study

• Examined four programs
  – Gator, Cydoor, SaveNow, and eZula
  – piggyback installation, adware and HTML rewriting
• Derived network signatures
  – look for the spyware “phoning home”
    – e.g., Gator traffic contains Gator/x.xx UserAgent and is sent to a *.gator.com host
  – signatures permit passive network monitoring

Method

• Network monitor deployed at UW
  – sniffs packets sent between UW hosts and Internet
  – gathered a 7 day Web trace
    • Aug. 26th – Sept. 2nd, 2003
  – looked for packets that match signatures
    • traffic matches signature ⇒ sender has spyware

The major result

<table>
<thead>
<tr>
<th></th>
<th>WWW</th>
<th>Gator</th>
<th>Cydoor</th>
<th>SaveNow</th>
<th>eZula</th>
</tr>
</thead>
<tbody>
<tr>
<td># clients</td>
<td>31,303</td>
<td>1,077</td>
<td>399</td>
<td>406</td>
<td>63</td>
</tr>
<tr>
<td>(% clients)</td>
<td>(100%)</td>
<td>(3.4%)</td>
<td>(1.3%)</td>
<td>(1.3%)</td>
<td>(0.2%)</td>
</tr>
</tbody>
</table>

• 5.1% of UW hosts have ≥ 1 of these programs
• This may appear small, but:
  – Only considers 4 spyware programs out of thousands
  – University may be non-representative
    • modem pool has 2.5x higher infection rate
  – a Gator vulnerability ⇒ 1000+ UW hosts at risk

Security flaws

• Gator & eZula “auto-update” their code, data
  – periodically download ZIP file, unzip into filesystem
• No integrity / authenticity checks on updates
  – Could attack with DNS spoofing or TCP hijacking
  – We could install an executable in “Startup” directory
  – Tens of millions of hosts susceptible
• They communicated flaws to both companies
  – Gator flaw was quickly repaired
Is there “at-risk” behavior?

- **# of web objects downloaded**
  - fewer than 1000 requests per week: 1.8% have Gator
  - more than 12,000 requests per week: 8.9% have Gator

- **# of executables downloaded**
  - none downloaded over week: 0.9% have Gator
  - one or more over week: 8.4% have Gator

- **using the Kazaa P2P client**
  - issued one or more Kazaa request: 38% have Gator
  - but...62% of spyware infections are in hosts that didn’t issue a Kazaa request

Two research studies

- **November 2003 study of adware within UW**
  - passive network measurement of entire campus
  - measured spread of four adware programs

- **Sneak preview of crawler-based study**
  - active retrieval of content from the Web
  - how much is “out there,” and who is spreading it

Two studies

- **November 2003 study of adware within UW**
  - passive network measurement of entire campus
  - measured spread of four adware programs

- **Sneak preview of crawler-based study**
  - active retrieval and analysis of Web content
  - how much is “out there,” and where is it coming from?

Method

- **Crawl subsets of Internet to find spyware**
  - used “heritrix” public domain crawler
  - downloaded .zip, .exe, .cab, etc. (programs)

- **Cluster of virtual machines to analyze programs**
  - “forked” a clean Windows VM per program
  - installed program, ran anti-spyware tool to analyze
  - O(1 min) per program
  - on 10-node cluster, O(15,000) programs per day
  - many performance optimizations possible

Major result

- **Web sites crawled: 12,000**
  - URLs retrieved: 23,714,927

- **# of executable files downloaded: 9,330**
  - # infected with spyware: 766 (8.21%)
  - unique spyware programs found: 137

 1 in 12 executables on the Internet have spyware!

What kind of spyware is out there?

<table>
<thead>
<tr>
<th>Behavior</th>
<th>% spyware</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adware</td>
<td>57%</td>
</tr>
<tr>
<td>Browser hijackers</td>
<td>56%</td>
</tr>
<tr>
<td>Keyloggers</td>
<td>0.06%</td>
</tr>
<tr>
<td>Dialers</td>
<td>0.1%</td>
</tr>
<tr>
<td>Backdoors / monitors</td>
<td>15%</td>
</tr>
</tbody>
</table>

- **Other stats:**
  - 58% try to evade discovery or removal
  - 32% monitor Web browsing behavior
  - most popular: eZula, 180 solutions, SaveNow
Where does it come from?

- spyware purveyors “troll” popular destinations
- blacklists aren’t terribly useful

Spyware Wrap-up

- Spyware affects many people
  - 5% of UW computers have adware
    - substantially underestimates all spyware
  - 1 in 12 executables on the Internet have spyware
- Most spyware appears benign
  - adware is the most rampant
  - but, trojans and monitors are on the rise
- Even “benign” spyware can harm you
  - hidden risk of security flaws, instability
  - no opportunity to mitigate or isolate