If we control these …

... we can monitor these
Types of Storm C&C Messages

- Activation (report from bot to botmaster)
- Email address harvests
- Spamming instructions
- Delivery reports
- DDoS instructions
- FastFlux instructions
- HTTP proxy instructions
- Sniffed passwords report
- IFRAME injection/report
Spam campaign mechanics

Botmaster

HTTP proxies

HTTP

Proxy bots

TCP

Workers
Campaign mechanics: harvest
Campaign mechanics: spamming

Botmaster

HTTP proxies

HTTP

Proxy bots

TCP

Workers
<table>
<thead>
<tr>
<th>MACRO</th>
<th>SEEN LIVE</th>
<th>FUNCTIONALITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0)</td>
<td>✓</td>
<td>Spam target email address.</td>
</tr>
<tr>
<td>(A)</td>
<td>✓</td>
<td>FQDN of sending bot, as reported to the bot as part of the preceding C&amp;C exchange.</td>
</tr>
<tr>
<td>(B)</td>
<td>✓</td>
<td>Creates content-boundary strings for multi-part messages.</td>
</tr>
<tr>
<td>(Cnum)</td>
<td>✓</td>
<td>Labels a field’s resulting content, so it can be used elsewhere through (V); see below.</td>
</tr>
<tr>
<td>(D)</td>
<td>✓</td>
<td>Date and time, formatted per RFC 2822.</td>
</tr>
<tr>
<td>(E)</td>
<td>✓</td>
<td>ROT-3-encodes the target email address.</td>
</tr>
<tr>
<td>(Fstring)</td>
<td>✓</td>
<td>Random value from the dictionary named string.</td>
</tr>
<tr>
<td>(Gstring)</td>
<td>✓</td>
<td>Line-wrap string into 72 characters per line.</td>
</tr>
<tr>
<td>(Hstring)</td>
<td>✓</td>
<td>Defines hidden text snippets with substitutions, for use in HTML- and plain-text parts.</td>
</tr>
<tr>
<td>(I)</td>
<td>✓</td>
<td>Random number between 1 and 255, used to generate fake IP addresses.</td>
</tr>
<tr>
<td>(Jstring)</td>
<td>✓</td>
<td>Produces quoted-printable “=&quot;20” linewrapping.</td>
</tr>
<tr>
<td>(K)</td>
<td></td>
<td>IP address of SMTP client.</td>
</tr>
<tr>
<td>(M)</td>
<td>✓</td>
<td>6-character string compatible with Exim’s message identifiers (keyed on time).</td>
</tr>
<tr>
<td>(N)</td>
<td></td>
<td>16-bit prefix of SMTP client’s IP address.</td>
</tr>
<tr>
<td>(Ostring: num)</td>
<td>✓</td>
<td>Randomized message identifier element compatible with Microsoft SMTPSVC.</td>
</tr>
<tr>
<td>(Pnum1[-num2]: string)</td>
<td>✓</td>
<td>Random string of num1 (up to num2, if provided) characters taken from string.</td>
</tr>
<tr>
<td>(Qstring)</td>
<td>✓</td>
<td>Quoted-printable “=&quot; linewrapping.</td>
</tr>
<tr>
<td>(Rnum1-num2)</td>
<td>✓</td>
<td>Random number between num1 and num2. Note, special-cased when used with (D).</td>
</tr>
<tr>
<td>(Ustring)</td>
<td></td>
<td>Randomized percent-encoding of string.</td>
</tr>
<tr>
<td>(Vnum)</td>
<td>✓</td>
<td>Inserts the value of the field identified by (Cnum).</td>
</tr>
<tr>
<td>(W)</td>
<td></td>
<td>Time and date as plain numbers, e.g. “20080225190434”.</td>
</tr>
<tr>
<td>(X)</td>
<td></td>
<td>Previously selected member of the “names” dictionary.</td>
</tr>
<tr>
<td>(Ynum)</td>
<td>✓</td>
<td>8-character alphanumeric string, compatible with Sendmail message identifiers.</td>
</tr>
<tr>
<td>(Z)</td>
<td>✓</td>
<td>Another Sendmail-compatible generator for message identifiers.</td>
</tr>
</tbody>
</table>

Table 2: Storm’s spam-generation templating language.
Figure 2: Snippet of a spam template, showing the transformation of an email header from template (top) to resulting content (bottom). The >-symbol indicates line continuations. Bold text corresponds to the formatting macros and their evaluation.
Campaign mechanics: spamming

- Botmaster
- HTTP proxies
- HTTP
- Proxy bots
- TCP
- Workers
- Spam cans
<table>
<thead>
<tr>
<th>CLASS</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Money mule scam</td>
<td>Attemps to enroll the victim in money laundering schemes</td>
</tr>
<tr>
<td>Personal ad scam</td>
<td>Fake dating/matchmaking invitations intended to convince victim to advance money</td>
</tr>
<tr>
<td>Job ads</td>
<td>Variant of money-mule scams, new “employee” is asked to forward money or goods</td>
</tr>
<tr>
<td>Self-propagation</td>
<td>Tricks or lures victims into executing malicious binaries¹</td>
</tr>
<tr>
<td>Phishing</td>
<td>Entices victims to enter sensitive information at fake bank sites or similars</td>
</tr>
<tr>
<td>Pharmaceutical</td>
<td>Pointers to web sites selling Viagra, Cialis, and other “male enhancement” products</td>
</tr>
<tr>
<td>Stock scam</td>
<td>Tries to convince victim to buy a particular stock supposedly about to increase in value</td>
</tr>
<tr>
<td>Other ads</td>
<td>Other kinds of advertising</td>
</tr>
<tr>
<td>Image spam</td>
<td>Image-based spam²</td>
</tr>
<tr>
<td>Other</td>
<td>Broken or empty templates, noise-only templates, etc.³</td>
</tr>
</tbody>
</table>

Table 3: Meanings of campaign classes.
Figure 5: Classes and instances of spamming campaigns identified over time.
<table>
<thead>
<tr>
<th>Self-propagation</th>
<th>Pharmacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>hotmail.com</td>
<td>8.24</td>
</tr>
<tr>
<td>yahoo.com</td>
<td>4.96</td>
</tr>
<tr>
<td>gmail.com</td>
<td>3.22</td>
</tr>
<tr>
<td>aol.com</td>
<td>2.40</td>
</tr>
<tr>
<td>yahoo.co.in</td>
<td>1.14</td>
</tr>
<tr>
<td>sbcglobal.net</td>
<td>0.97</td>
</tr>
<tr>
<td>mail.ru</td>
<td>0.82</td>
</tr>
<tr>
<td>shaw.ca</td>
<td>0.64</td>
</tr>
<tr>
<td>wanadoo.fr</td>
<td>0.63</td>
</tr>
<tr>
<td>msa.hinet.net</td>
<td>0.60</td>
</tr>
<tr>
<td>msn.com</td>
<td>0.58</td>
</tr>
<tr>
<td>excite.com</td>
<td>0.49</td>
</tr>
<tr>
<td>yahoo.co.uk</td>
<td>0.43</td>
</tr>
<tr>
<td>rediffmail.com</td>
<td>0.34</td>
</tr>
<tr>
<td>comcast.net</td>
<td>0.32</td>
</tr>
<tr>
<td>ig.com.br</td>
<td>0.31</td>
</tr>
<tr>
<td>verizon.net</td>
<td>0.27</td>
</tr>
<tr>
<td>earthlink.net</td>
<td>0.27</td>
</tr>
<tr>
<td>btinternet.com</td>
<td>0.26</td>
</tr>
<tr>
<td>t-online.de</td>
<td>0.25</td>
</tr>
</tbody>
</table>
Campaign mechanics: reporting

- **Botmaster**
- **HTTP proxies**
- **HTTP**
- **Proxy bots**
- **TCP**
- **Workers**
Measurements: delivery efficacy

![Bar chart showing occurrence in % of different error types: SMTP 220, RCPT, Success, MAIL, MX, TCP, DATA, HELO. The chart indicates the highest occurrence is for SMTP 220, followed by RCPT, Success, and then MAIL and MX, with TCP, DATA, and HELO having the lowest occurrences.]