Shawn Roberts, Fotolia

Using tarpits to trap spammers

TAR TANGLE

A tarpit slows down spam by reducing the effectiveness of the spammer's weapons. BY TOBIAS EGGENDORFER

Many email addresses appear on the web, and spammers employ harvester applications to collect those addresses for future mailings. Some victims fight back with a tool called a tarpit. A tarpit is an automatically-generated website that baits a harvester with a complex tangle of meaningless URLs (Figure 1).

The longer you can pull the wool over the harvester's eyes, the longer the list of tarpitted addresses will become. Under ideal circumstances, the harvester would end up with a list full of tarpitted links.

To allow this to happen, the bait site must publish more links to itself than an average website. I ran a mini spider I developed myself against 23,000 pages and determined that each page had an average of 6.4 new links. The typical tarpit may publish 20 new links per page – three times the average value. Every round adds to the number of tarpitted links in the harvester's list of sites.

The only way for a harvester to avoid the trap is to limit the amount of time spent in each domain, but setting a time limit causes the harvester to stop before it is finished, which means the spammer does not succeed in harvesting all the addresses on your site. You could say that the more website operators who set up tarpits, the more difficult life becomes for spammers.

Practical trials with the tarpit program show that it can fool harvesters for days, thus proving the effectiveness of the HTTP tarpit. To make the tarpit even more effective, some sites add sophisticated camouflage or set up tarpits with links to other tarpits.

Speed Spamming

Spam is only worthwhile if a spammer can deliver a huge volume of mail in a short period of time with minimal overhead. Per-message charges have actually been suggested at various times as a means of combating spam, however, many experts worry that the virtual stamp charges would also affect legitimate mail users.

Slowing down spam relays makes spamming computationally expensive. In an ideal case, connections would be slowed to a point where they become unattractive for the spammer. And this is exactly what SMTP tarpits attempt to do. Various types of SMTP tarpits address different scenarios. Some SMTP tarpits are designed to protect a specific site, whereas others aim to completely remove spam from the Internet.

This group of heroes exists only for the purpose of accepting mail from spammers. The idea is to point the MX record in a domain at the SMTP tarpit and to publish a huge volume of email addresses that nobody uses. If everything works to plan, every single mail message sent to these addresses should be spam.

Improving the World

Stand-alone SMTP tarpits such as SMtarpit [2] are designed to free the Internet from spammers. In theory, if a spammer stumbles into an SMTP tarpit, they are

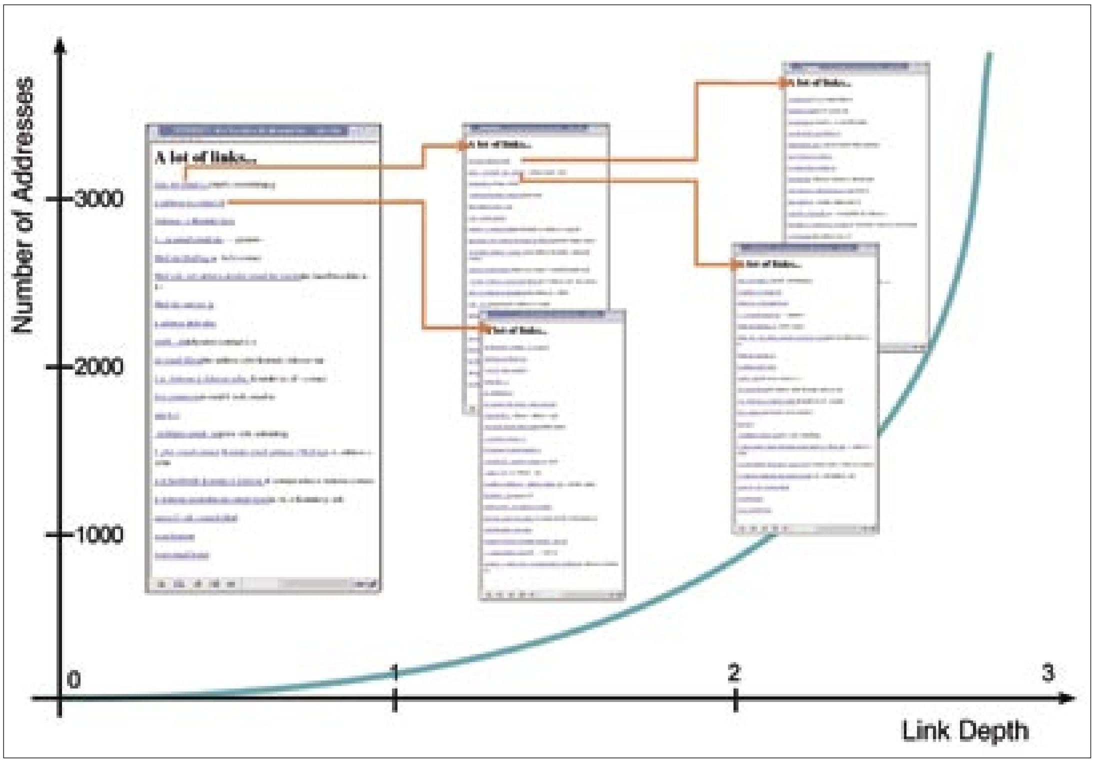

Figure 1: HTTP tarpits are designed to trick address collection tools that grab email data off websites. Well-prepared traps can lead harvesters into a tangle of unproductive links.

likely to get trapped. Unfortunately, modern bulk mailers are multi-threading or multi-tasking applications. They open multiple parallel connections to mail servers. If one of the targets is a tarpit, the bulk mailer loses one connection, however, this does not typically have much effect, as a client can open a connection to an SMTP server from any high port number. There are more than 64,000 high ports, and the bulk mailer can send any number of messages over each connection.

Thus, each tarpit can block one of 64,000 possible connections. The spammer can still misuse all the other mail servers, and the damage that is done to the spammer is negligible. The argument that the proponents put forward, that many distributed SMTP tarpits would do the trick, is easy to disprove.

In fact, to block 25 percent of a bulk mailer's connections, 25 percent of all mail servers would have to be tarpits. Assuming estimated figures of 25 million mail servers world wide, this would mean having 7.5 million tarpits, which is clearly an unrealistic figure.

Spam-Blocking Tools

Most bulk mailers now use aggressive timeouts and tend to interrupt connections as early as possible. These tactics make stand-alone SMTP tarpits relatively ineffective.

The fact that bulk mailers try to protect themselves does have a useful side effect: the second category of SMTP tarpits acts as an SMTP proxy for an existing mail server. The tarpit accepts incoming connections addressed to the MTA and slows them down.

Some elegant tools in this category are OpenBSD *spamd* [3] and Lutz Donnerhacke's SMTP-Wrapper, which unfortunately uses static blocking lists. The OpenBSD tool *spamd* has a black list of known spam IPs and a whitelist of IP addresses of known, legitimate mail servers. All others are handled by *spamd*'s graylisting mechanism, which is used to populate the whitelist.

To do so, the daemon responds with a temporarily unavailable error code; if the sender does not retry, the tool assumes that the IP belongs to a spammer and adds the address to the blacklist. Senders who do retry are probably legitimate and are added to the whitelist.

Whenever a message from an address that is identified as belonging to a spammer reaches the program, the program will artificially slow down the connection. As the bulk mailer uses this characteristic to detect a tarpit, the mailer will drop the connection as quickly as possible. Dropping the connection prevents the spammer from delivering junk mail to the local mail server.

spamd can also slow down connections from senders who are recorded on the graylist.

Again, spammers will tend to give up rather quickly when handled like this. David Purdue [3] reports that the decision to stop is made within a couple of seconds of the connection being established.

Thus, a tarpit can effectively prevent most spam from ever-reaching the local mail server because bulk mailers usually will tend to drop the connection.

This protects the server operator against spam, while preserving bandwidth that the spammer would otherwise misappropriate.

Conclusions

An HTTP tarpit is a useful preventive measure that is more than capable of annoying email harvesters. SMTP tarpits are useful, but not as a means of fighting back against spammers.

To fight back against spammers, you might want to consider setting up an SMTP tarpit as a proxy for your own MTA. Due to aggressive timeout policies, bulk mailers will then stop trying to deliver spam quickly. ■

INFO

[1] Giovanni Donelli, "Email Interferometry": Spam Conference 2006, MIT, Cambridge, MA

[2] Paul Grosse, SMtarpit v0.6.0: *http://www.fresh.files2.serveftp.net/smtarpit/*

[3] David Purdue, "Adventures in the tar pit – Implementing OpenBSD's spamd": AUUG 2005, Sydney and *http://www.openbsd.org/spamd/*