Evaluating The Use of Spam-triggered TCP/IP Rate Control To Protect SMTP Servers

Minh N. Tran
Grenville J. Armitage

Talk outline

- **What is Spam?**
- Our anti-spam tool: MT Proxy
- A basic email architecture
- Traditional and new anti-spam techniques
- Design and functionality of MT Proxy
- Experimental validation
- Research limitation and future work
- Conclusion

What is Spam?

- "Mass unsolicited electronic mail"

  (RFC 2505, "Anti-Spam Recommendations for SMTP MTAs ",
  http://www.ietf.org/rfc/rfc2505.txt)

- "Unsolicited Bulk Email ("UBE")"

  - Unsolicited: "Recipient has not granted verifiable permission for the message to be sent"
  - Bulk: "Message is sent as part of a larger collection of messages, all having substantively identical content"

  (The Spamhause project, "Definition of spam",
  http://www.spamhaus.org/definition.html)

- "Unsolicited e-mail, often of a commercial nature, sent indiscriminately to multiple mailing lists, individuals, or newsgroups; junk e-mail"

  (Dictionary.com, http://dictionary.reference.com/search?q=spam)

Talk outline

- What is Spam?
- Our anti-spam tool: MT Proxy
- A basic email architecture
- Traditional and new anti-spam techniques
- Design and functionality of MT Proxy
- Experimental validation
- Research limitation and future work
- Conclusion

Our anti-spam tool: MT Proxy

- A new anti-spam approach
- Inspired by Marty Lamb's exhortation
  - "We want to cause spammer pain"
    (Marty Lamb, "Using Statistic to cause Spammers Pain", February 2003, http://www.martiansoftware.com/articles/spammerpain.html)
- Two goals of MT Proxy:
  - To cause resource consumption at spammer end
  - To avoid negative impact of legitimate emails being misclassified as spam
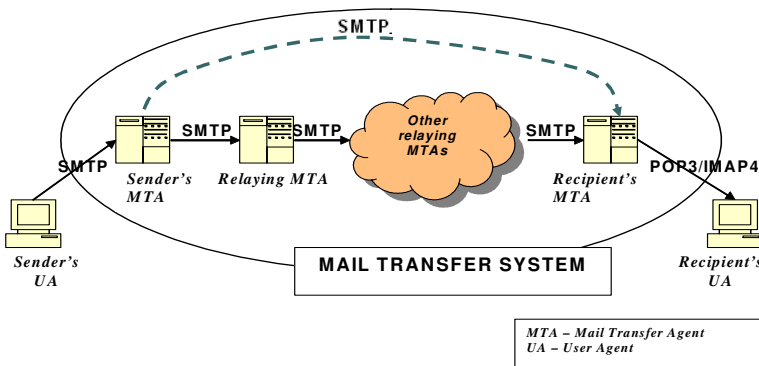
Talk outline

- What is Spam?
- Our anti-spam tool: MT Proxy
- **A basic email architecture**
- Traditional and new anti-spam techniques
- Design and functionality of MT Proxy
- Experimental validation
- Research limitation and future work
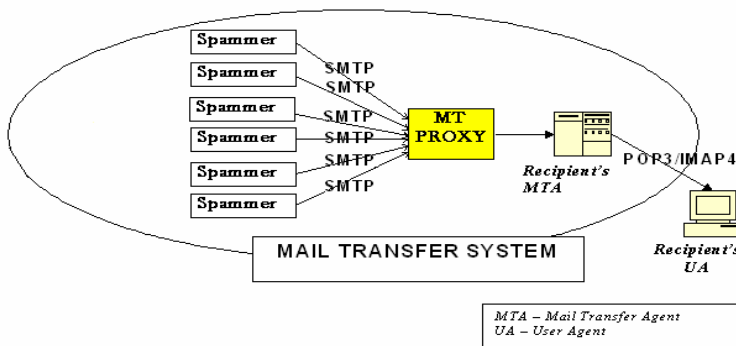- Conclusion

A basic model of email transfer

- Simple Mail Transfer Protocol (SMTP, defined in RFC 821)

Role of MT Proxy

- MT Proxy sits in front of recipient's MTA to intercept SMTP traffic coming to this server

Where MT Proxy sits in the Email Transferring System

Talk outline

- What is Spam?
- Our anti-spam tool: MT Proxy
- A basic email architecture
- **Traditional and new anti-spam techniques**
- Design and functionality of MT Proxy
- Experimental validation
- Research limitation and future work
- Conclusion

Traditional anti-spam techniques

- White and black listing:
  - Focus on addresses associated with the email
  - Emails in White list are legitimate
  - Emails in Black list are spam
  - Challenge response mechanism
- Rule-based filtering:
  - Inspect the actual content of the email
  - Bayesian algorithm: most spam events are dependent
    - Paul Graham method (P. Graham, " A Plan for Spam", August 2002, http://www.paulgraham.com/spam.html)
    - Gary Robinson method (Robinson G. ,"Spam Detection", October 2003, http://radio.weblogs.com/0101454/stories/2002/09/16/spamDetection.html)

Challenges for traditional methods

- Trade-off between False Positive and False Negative
  - False Positive:

    A legitimate email is incorrectly identified as spam
  - False Negative:

    A spam email is incorrectly classified as non-spam
  - ISPs implement more aggressive anti-spam methods
    - Loss of over $50 per person per year and $3.5 billions per a U.S. business in 2003 due to false positive (Research by Ferris Inc.)
  - Need to eliminate false positives
- Traditional methods leave no painful impact on spammer

ATNAC2004   {mtran, garmitage}@swin.edu.au   http://caia.swin.edu.au   12/11/2004   Page 11

New anti-spam approaches

- Sender Policy Framework (SPF)
  - Domain authentication technique to identify spam forgery
  - Maintain registered domain names and their associated mail servers
- Anti-spam router (ASR) of TurnTide
  - Allocate different Quality of Service (QoS) for different incoming email traffic according to its spam level
- Microsoft's "stamp of approval"
  - Delay is added to SMTP traffic through cryptographic puzzles solved by sender
- MT Proxy
  - Eliminate the negative consequences of false positives
  - Shift back the cost to spammers
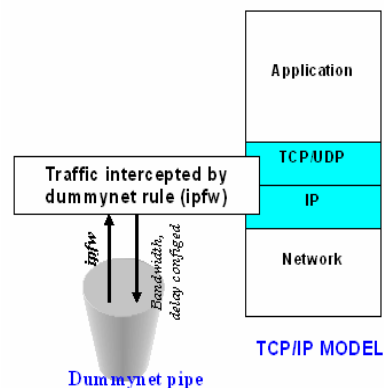
ATNAC2004   {mtran, garmitage}@swin.edu.au   http://caia.swin.edu.au   12/11/2004   Page 12

Swinburne University of Technology   6

Talk outline

- What is Spam?
- Our anti-spam tool: MT Proxy
- A basic email architecture
- Traditional and new anti-spam techniques
- **Design and functionality of MT Proxy**
- Experimental validation
- Research limitation and future work
- Conclusion

Design and functionality

- FreeBSD proxy server intercepting email traffic
- Use blacklisting and content filtering
- Blacklisting: local and Internet DNS server
- Content filtering: spammer experiences slower connection in real-time
- Traffic is shaped at TCP/IP level using FreeBSD kernel's resident: ipfw and dummynet
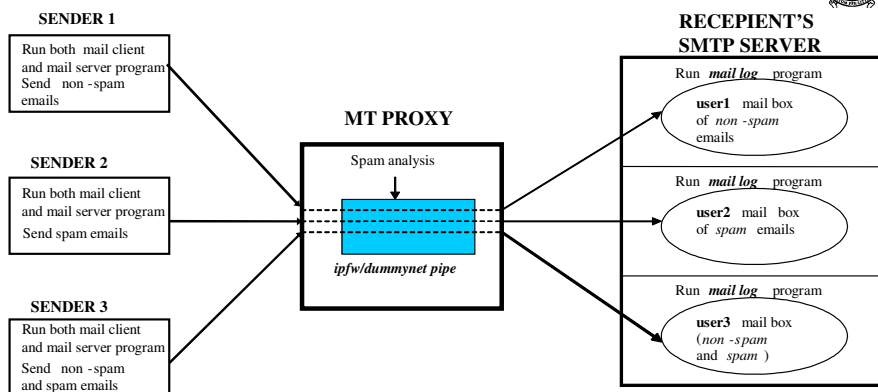
Talk outline

- What is Spam?
- Our anti-spam tool: MT Proxy
- A basic email architecture
- Traditional and new anti-spam techniques
- Design and functionality of MT Proxy
- **Experimental validation**
- Research limitation and future work
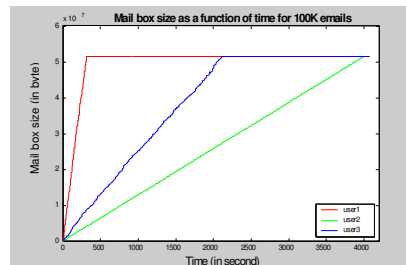- Conclusion

---

Experimental validation

SENDER 1

Run both mail client and mail server program. Send non-spam emails

SENDER 2

Run both mail client and mail server program. Send spam emails

SENDER 3

Run both mail client and mail server program. Send non-spam and spam emails

MT PROXY

Spam analysis

ipfw/dummynet pipe

RECEPIENT'S SMTP SERVER

Run *mail log* program

user1 mail box of *non-spam* emails

Run *mail log* program

user2 mail box of *spam* emails

Run *mail log* program

user3 mail box (*non-spam* and *spam*)

Testbed for evaluating the efficacy of MT Proxy

Swinburne University of Technology

Experimental validation: Spam vs non-spam

- Clients use open source software smtpclient to send 500 non-spam,spam,non-spam/spam emails to user1, user2, user3 respectively
- Each client send emails as fast as possible, but do not overlap itself
- From the graph, MT Proxy has proved to be capable of slowing down spam

Mail box size as a function of time for 100K emails

Experimental validation: Email size

- When email size increases, the relative time delay between spam and non-spam also increases
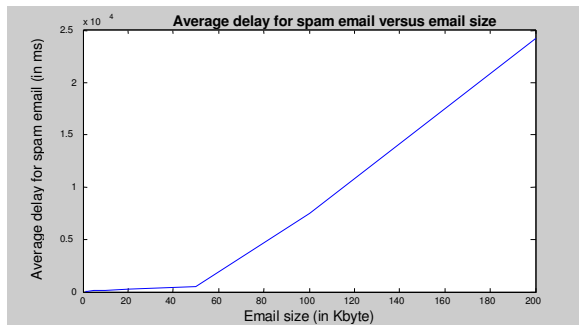- When email size increases, the mail box size reduction also improves

Email size (in Kbyte)	Average delay for spam email (in ms)	Number of non-spam emails received per second	Number of spam emails received per second	Spam mail box size reduction per second
1	1	6.59	6.45	143
2	54	4.84	4.32	1065
5	108	2.60	2.06	2813
10	120	2.50	2.04	4710
20	240	2.40	1.54	17613
50	450	2.11	1.08	52736
100	7500	1.57	0.13	147456
200	24200	1.00	0.04	196608

Spam reduction for different email sizes (100% spam case)

Experimental validation: Email size

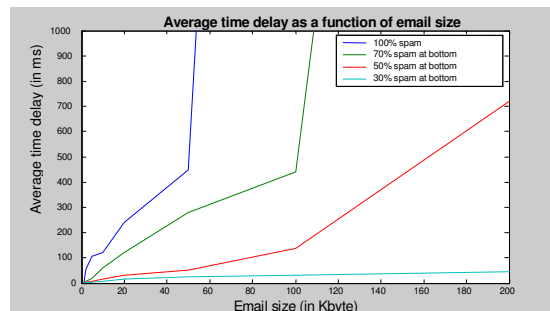• After the threshold, spam emails are received at a significant time after non-spam emails

Average delay for spam email versus email size

Time delay as a function of email size

Experimental validation: Spam structure

• 4 types of email structure in which the top part is non-spam and the bottom part is spam
• The earlier spam appears in the email, the more effective MT Proxy is

Average time delay as a function of email size

- 100% spam
- 70% spam at bottom
- 50% spam at bottom
- 30% spam at bottom

Average time delay as a function of email size for 4 types

Talk outline

- What is Spam?
- Our anti-spam tool: MT Proxy
- A basic email architecture
- Traditional and new anti-spam techniques
- Design and functionality of MT Proxy
- Experimental validation
- **Research limitation and future work**
- Conclusion

ATNAC2004 {mtran, garmitage}@swin.edu.au http://caia.swin.edu.au 12/11/2004 Page 21

Research limitation and future work

- The architecture of MT Proxy does not work very effectively for spam emails with small size
  - Email size ranges from 1Kbytes to 11Kbytes with mean of 4.64 Kbytes (6955 spam emails logged by our University's IT department on 2 Aug 2004)
- MT Proxy is more effective against email bodies where signs of spam occur early
- Improved version of MT Proxy learns which source IP addresses had attempted to send spam to rate-limit their subsequent connection
- Sending many small emails will be equivalent to sending a single large consecutive email

ATNAC2004 {mtran, garmitage}@swin.edu.au http://caia.swin.edu.au 12/11/2004 Page 22

Talk outline

- What is Spam?
- Our anti-spam tool: MT Proxy
- A basic email architecture
- Traditional and new anti-spam techniques
- Design and functionality of MT Proxy
- Experimental validation
- Research limitation and future work
- **Conclusion**

Conclusion

- MT Proxy sits in front of recipient's SMTP server
- MT Proxy has proved to have certain contribution to the arsenal of anti-spam techniques
- MT Proxy can effectively slow down traffic from spammers using FreeBSD's kernel resident ipfw/dummynet
- MT Proxy's approach also avoids the damaging consequences of false positives because all email eventually gets through
- We are developing an improved version of MT Proxy with long term memory of who is sending spam