

Email Security Techniques

David Gibbs

david@midrange.com / dgibbs@ptc.com

With respect to all session attendees
and the presenter,

please turn Cell Phones off
or place on vibrate
during class sessions!



Agenda

- **What is Spam?**
- **Defend Your Company.**
- **Protect Your Company.**
- **What does midrange.com do?**
- **Resources**

What Is Spam?

- Officially known as “Unsolicited Commercial Email”
- Basically: Commercial email that you didn’t request and don’t want.
- According to the European Network and Information Security Agency, spam now accounts for more than 95% of mail traffic.
- It’s a drain on your company’s resources and a waste of your employees’ time.

Why Send Spam?

- It makes money -- sadly, people fall for it.
- It's cheap -- spammers rarely use their own resources to send messages.
- It's an easy way to get personal information.



Who Sends Spam?

- Nefarious ...
 - Purveyors of “products”: ‘buy this genuine, imitation, Rolex watch’ or various & sundry ‘marital’ aids
 - Scammers: ‘help me transfer 7.8 trillion US Dollars from my dead Nigerian husband’s account.’
 - Phishers: “This is your bank, please confirm your social security number and account number so we can release your funds.’
- Not so nefarious ...
 - Vendors: ‘visit us at the COMMON expo in booth 123’.
 - Marketers: ‘buy my computer product, it’s better than the one you’re using’





Identify Spam

- “This is not spam” (yes, this is first on the list for a good reason)
- Asks for your password, credit card number, social security number, etc.
- An online service tells you to download a program to reset your password.
- Appears to come from your bank. Always look twice if you get email from a bank.
- From PayPal, eBay, amazon.com, or other major online merchants.
- Webmail system tells you that you are over quota (especially if you don't use that webmail system).
- Any unexpected message from ‘support@yourdomain.com’ (especially since we're the ones who would be sending that kind of message anyway).

How Is Spam Sent?

- Open relays (mail servers that allow mail to be relayed through them without any authentication)
- Zombie computers – Also known as ‘bots. These are computers infected with a virus that gives someone else control (often referred to as a ‘bot herder’).
- Free email providers like AOL, Gmail, Hotmail, etc.

Defend Your Company

Block Lists

- DNS Block Lists (DNSBL) are community maintained lists of known spam sources.
- Many mail servers can be configured to automatically block mail from servers listed.
- Pros
 - Able to block much of the spam in the world.
 - Community maintained, very fast to react.
- Cons
 - Sometimes too aggressive.
 - Community maintained ... and may go down.

Block Lists -- Advice

- Pick your block lists carefully.
 - Some are too aggressive (spamcop.net).
 - Some list servers that do not necessarily send spam (example: rfc-ignorant.org).
 - Some make it excessively difficult to get a IP address removed.
- Monitor the block lists (if they go down, sometimes they return positive results for all queries)

Sender Policy Framework

- SPF : Sender Policy Framework
- Is the message from a specific domain sent from the right server?
- Uses DNS 'txt' records to list all the servers that a domain uses to send mail.
- Pros: Can be used to verify a message from 'someone@example.com' actually came from a server that 'example.com' uses.
- Cons: Not many organizations use SPF. Additionally, some spammers have figured out how to setup SPF quickly.



Signature Checking

- Used to verify that a message came from the organization it appears to be from and that it has not been tampered with en-route.
- Messages are signed with public key encryption techniques that can be used to verify a message is legitimate.
- How it works
 - Before a message is delivered, a special header is added that contains a digital signature based on content.
 - The receiving server looks up the public key of the domain the message appears to be from.
 - The public key is used to decrypt the signature and verify the message.
- Pros: Very secure, almost impossible to forge.
- Cons: Not all organizations use mail signatures.



Signature Checking

- Implementations
 - DKIM -- DomainKeys Identified Mail
 - DomainKeys (deprecated, but still used by some major mail providers)
 - Both DKIM and DomainKeys were originally championed by Yahoo.

Sidebar: Why DNS?

- DNS is a very lightweight protocol.
- Queries to DNS return very little data and are very fast.

White List

- A list of email addresses that are known not to send spam.
- The list should be automatically maintained by your anti-spam software.
- Keep in mind: Spammers usually spoof the FROM address.
- Don't make a decision JUST based on a black list or white list.



Grey Listing

- When mail comes in from an unknown server, the message is temporarily rejected with a 'soft' failure.
- If the message is redelivered before the greylisting duration has expired, it is rejected again with a soft failure.
- If the message is redelivered after the greylisting duration has expired, it is accepted and future deliveries from that server are not delayed.
- This is implemented on the mail server or possibly with a plug-in.

Grey Listing

- Rationale
 - Normal mail servers will retry a mail delivery after a reasonable period of time.
 - Spammers don't care if individual messages are actually delivered, and won't retry the delivery.
- Risks
 - Mail from new senders WILL be delayed.
 - Some mis-configured mail servers handle a temporary failure message as a permanent failure. Although rare, some major mail servers have never been fixed.

No Listing

- Create a backup MX that deliberately rejects messages with a soft failure.
- Since spammers often target the lower priority MX records, they will try to deliver mail to that server.
- Legitimate mail servers will see the soft failure and retry the delivery later.

Rate Limiting

- Also known as throttling.
- Impose a limit on frequency of SMTP commands after a certain number of bad addresses presented.
- Doesn't really reduce the amount of spam, but can lighten the load on your mail server.

Content Filters

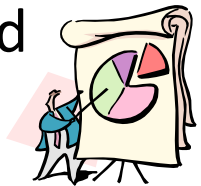
Pattern Match

- Matching on words and patterns.
- Patterns should be complex, not simple.
- Should be used along with other methods.
- Don't reject a message simply because of the presence of a specific word.
- Example pattern filters:
 - “(wife | son | brother | daughter) of the late”
 - “(my name is | i am) (mrs | engr | barrister | dr | prince)”
 - “give you (fund | money | total | sum | contact | percent)”

Content Filters

Bayesian Analysis

- Bayesian analysis calculates the probability of a message being spam by looking at its contents. Unlike content filters, Bayesian analysis learns from the spam and non-spam. This results in a very adaptive and efficient filtering technique.
- Is probably one of the most effective spam detection methods available.
- The Bayesian analysis software analyzes email content and evaluates the frequency of words.
- If certain words are encountered that are often associated with spam, the word is classified as 'spammy'.
- Conversely, if certain words are encountered that are rarely associated with spam, the word is classified as 'not spammy' (also known as "Hammy").
- Bayesian software remembers the words it processes and keeps a score.



- Bayesian analysis software must be trained.
 - Both spam and non-spam
 - Requires several hundred messages to be effective.
- Neutral words ('a', 'is', 'some', etc) are ignored.
- It continues to learn.
 - If a new word is encountered, and the other words in the message are associated with spam, that word will be associated with spam.
 - If an existing word is encountered that is associated with spam, but the other words in the message are not associated with spam, the existing word will be less associated with spam.

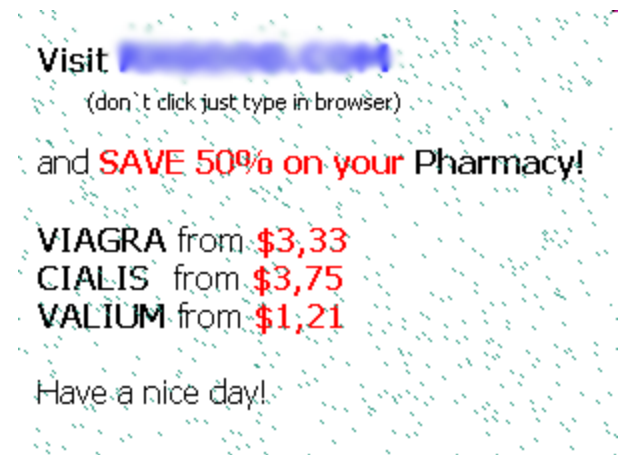
Content Filters

Bayesian Analysis

- Risks
 - Can make mistakes
 - Should be taught from mistakes.
 - Less likely to make mistakes with larger initial training.
 - Poison
 - Spammers try to poison Bayes engines by putting non-spam words in their messages to lower the spam score and raise the spam score of those non-spam words.
 - This is why you might see passages from classical novels in spam messages.
 - Good Bayesian implementations learn well enough where this doesn't cause a big problem.

Content Filters Bayesian Analysis

- Weakness – Image Spam
 - Bayes only works on text. Some spammers put their messages in image files.
 - There has been some effort to use OCR to analyze images to determine if they are spam.
 - Spammers have tried obfuscating the images to defeat OCR software.



Anti-Virus

- Many viruses are delivered via email.
- Mail should be scanned before it is completely accepted by the mail server.
- Any message that is found to be infected should be rejected immediately.
- Don't bother trying to notify the sender ... the 'FROM' address is almost always forged.
- Stay current on virus signatures.



Manage Your Backup Mail Server

- Many backup mail servers have less protection than the primary server.
- For this reason, some spammers target them (this is the rationale behind 'no listing').
- Ensure your backup server has the same protections that your primary server does.
- This includes knowing who your users are and copies of all spam rules.



Techniques NOT to use

- Simple content analysis.
 - Could easily block legitimate mail.
 - Some 'dirty' words are the unintended base for real words.
- Domain based white or black lists.
 - Spammers usually forge the 'from' address, so you could be allowing or blocking mail in error.
- Permission-based white lists (Challenge-response)
 - Response messages are sometimes interpreted as spam.
 - Sometimes the sender is not really a human and can't respond.
 - Innocent people will get the challenge, due to address forging, and have no clue as to why.

More Techniques NOT to use

- Blocking or accepting based on downstream headers
 - Headers are frequently forged
 - Sometimes legitimate mail is routed through unexpected servers (i.e., mailing list servers).
- Never make assumptions about the mail you will or won't receive.

Which Techniques To Use?

- Good mail systems use multiple techniques.
 - No one technique is sufficient.
 - Some things are mandatory (anti-virus)
 - Two simple examples:
 - A block list check may say the message is from an open relay, but Bayesian analysis says there is a low probability that the message is spam and it's virus free.
 - Sender isn't on a block list, no viruses, but Bayesian analysis says there is a high probability that the message is spam.

How To Implement?

- Build your own (my personal preference)
 - Gateway to & from an internal mail server.
 - Lots of open source software that can be used to build a very solid anti spam solution.
 - SpamAssassin
 - SpamBayes
 - Clam Antivirus
- Use a mail appliance or filtering service
 - Barracuda Spam Firewall
 - Postini (Google)
 - MessageLabs (Symantec)

Protect Your Company

(from being identified as a spammer)

Keep Your House In Order

- An integral part of any mail security plan is to make sure your own mail servers are behaving correctly.
- General Protections
 - The obvious ...
 - Workstation Anti-Virus.
 - Server Anti-Virus.
 - Malware checking.
 - Keep current on software.
 - Educate your users.
 - Not so obvious ...
 - Authenticated Relay -- Allows mail to be sent using your server only by verified users.
 - Secure Communication (SSL).
 - Protect email addresses on web sites.



Sign Outbound Mail

- DKIM
- DomainKeys
- Can use both.
- Requires that all mail go out through a central server (which isn't a bad thing).



Establish & Publish Policies

- Publish Sender Policy Framework (SPF) information.
- List your outbound mail servers in DNS records so recipients can verify a message came from the correct server.
- Also requires that all mail go out through a central server.



Prevent Unauthorized Mail

- Block port 25 for all systems except official mail servers.
- Route all outbound mail through corporate mail server.
- Use authenticated sending.
- Encrypt server-to-server communication with SSL (becoming much more common).



Check Your IP Addresses

- Check your server IP's to ensure they are not on block lists.
- Ensure your server is not an open relay.
- There are many resources on the web to do these checks.



Prevent Back Scatter

- Back scatter is what happens when a mail server rejects a message that used a forged FROM address.
- People often get meaningless virus and spam warnings because their email address was used in spam.
- Backscatter messages often include the original message ... which can be identified as spam.

Don't Be Bad

- Confirm mailing lists
 - Use “Closed Loop Authentication”
 - Buy from “Reputable” sources
 - Get a guarantee. Some mailing list sources say they are 100% opt-in, but rarely are.
- Don't harvest addresses
 - Very easy to detect (some people use single purpose email addresses for online forums).
 - Our community is (relatively) small.

Don't Be Bad

- Correct problems promptly
 - If someone asks to be removed from your mailing list, remove them immediately and confirm the removal. Give them a incident number if you use a tracking system.
 - Keep track of people who want to removed so you don't add them back in.
- Comply with the 'CAN-SPAM' laws.
 - Targeted at legitimate email marketing.
 - Has no impact on the true spammers.

Don't Be Bad – CAN SPAM

- Don't use false or misleading header information.
- Don't use deceptive subject lines.
- Identify the message as an ad.
- Tell recipients where you're located.
- Tell recipients how to opt out of receiving future email from you.
- Honor opt-out requests promptly.
- Monitor what others are doing on your behalf.

Sign up for Feedback Loops

- Feedback Loops (FBL) are services provided by major ISP's to tell mailers when their messages are reported as spam.
- Some FBL's are quite useful ...
 - Yahoo (tells you who reported your mail as spam, so you can directly address the issue).
- Others are a waste of time ...
 - AOL
 - Rackspace
 - Both AOL & Rackspace redact the email address of the person reporting spam, so there's no way to respond.
- Links for various FBL's can be found in the resources section.

What Does midrange.com Do?

- Inbound mail security
 - Block lists using specific providers to completely block known spam sources.
 - Grey list for 5 minutes
 - Anti-Virus (Clam anti-virus)
 - SpamAssassin (spam score > 5 is spam, > 10 is discarded)
- Outbound mail security
 - SPF
 - DKIM
 - DomainKeys
 - Anti-Virus
 - Registered with multiple FBL programs.

What Does midrange.com Do?

- Mailing lists
 - Closed loop subscription confirmation. This ensures nobody is involuntarily subscribed.
 - Email addresses on list archives are obfuscated.
 - Any system the messages are gated to must obfuscate email addresses.

Questions?



Resources

- Federal Trade Commission -- <http://www.ftc.gov/spam>
- Mail Servers (recommended)
 - Sendmail – <http://www.sendmail.org>
 - Postfix – <http://www.postfix.com>
- SpamAssassin (<http://spamassassin.apache.org>)
- Grey Listing (<http://www.greylist.org>)
- DSL BL
 - SpamHaus -- <http://www.spamhaus.org/>
 - Composite Block List -- <http://cbl.abuseat.org/>
 - SpamCop -- <http://www.spamcop.net/>
 - Not Just Another Bogus List -- <http://njabl.org/>
 - OpenRBL -- <http://openrbl.org/>

Resources

- SPF -- <http://www.openspf.org/>
- DomainKeys -- <http://domainkeys.sourceforge.net>
- DKIM -- <http://dkim.org/>
- SpamCop.net
- Project Honeypot -- <http://www.projecthoneypot.org>

Resources

- Feedback Loops
 - Yahoo -- <http://feedbackloop.yahoo.net>
 - Road Runner -- <http://feedback.postmaster.rr.com>
 - Comcast -- <http://feedback.comcast.net>
 - Cox -- <http://fbl.cox.net>
 - USA.net -- <http://fbl.usa.net>
 - MSN / Hotmail -- <https://postmaster.live.com/Services.aspx>
 - AOL -- <http://www.postmaster.aol.com/fbl> (don't bother, see below)
 - Rackspace -- <http://fbl.apps.rackspace.com/> (don't bother, below)
 - Earthlink – Send an email at fblrequest@abuse.earthlink.net and provide your IP range, domains, contact information, and the email address for sending complaints.
 - Outblaze – Send an email with a feedback loop request at postmaster@outblaze.com
- Most of the FBL services are actually run by ReturnPath ... so expect a marketing email when you sign up for any of the loops.

Creative Commons License

- This work is licensed under the Creative Commons Attribution-Share Alike 3.0 United States License.
- To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/3.0/us/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

