Global Information Assurance Certification Paper

How to fight SPAM and Phishing

GIAC Security Essentials
Certification (GSEC)
Practical Assignment
Version 1.4c

Option 1

By Daniel Arnanz

Table of Contents

Abstract/Summary

This paper is designed to give an idea of what SPAM and phising are, and how to protect us against these threats. SPAM is a problem that worries both companies and end users around the world, and from what we see it seems it is getting worse everyday. In this paper I will explain what the SPAM problem is today and what we can expect to see in the near future. We will see different methods and ideas of trying to tackle SPAM from the point of view of a company with corporate email accounts, also I will try to explain why SPAM exists and why it is good business for some people.
Phising is another email related problem that has been a huge concern especially for financial institutions and also for end users for a few years now.

An Introduction to SPAM and phishing

SPAM

Nowadays when you say the word SPAM, everyone knows that you are talking about mail delivered to your Internet mail account from someone that you don't know, normally offering you some kind of product/service of dubious reputation that you can buy in Internet visiting some link attached. Everyone knows this because the email spam problem is getting bigger and bigger, and nearly everyone receives some amount of SPAM everyday.
The reason why SPAM exists, in the first place, is obviously for spammers to make money. And it is very simple, they make money advertising things for someone else and earning some small percentage (lets say 1$) for every hit. If they send an email to 1 million people (this costs them practically nothing) offering some kind of product and just 10 of them actually buy it, the spammer has already made a 10$ profit. So if they keep doing it full time to a big number of email addresses they can make a living as a Spammer. Only if people stopped buying things from these companies that create spam, the problem would come to an end immediately, because if they didn't make a profit, then there is no reason to carry on sending emails. But far from this, the reality is that people still buy and will continue to buy from these places, so we have to find a way of stopping the spam by developing antispam technologies.
Because of this opportunity to make money, spammers are always trying to find new ways of spreading these offers to potential buyers and that is why we see spam being used in other environments like SPAM in Instant Massaging (SPIM), Spam in voice over IP (SPIT), web SPAM or Blog Spam.

Instant Messaging Spam (SPIM)

Instant messaging is known for being insecure, as information normally travels in clear text in Internet and because of the danger of virus-adware propagation

like the adware called *Buddylinks* or *Osama* that sends a message to everyone in you buddy list. But apart from this, IM is also used to send unwanted messages, instead of using traditional email, the '*spimmers*' send the messages to users of instant messaging programs. This is not new, but with the increase of IM users SPIM is becoming an increasing annoyance. SPIM can be even more annoying than SPAM because their messages appear as pop-ups, so it is more difficult to ignore them. Normally most of spim is primarily generated using bots and is related most often to pornographic matters.

A way to avoid spim would be to accept messages only from users who figure on our buddy list, but this can be easily bypassed using spyware tools that find out our buddy list. So by masquerading themselves as a buddy, spimmers could eventually be able to send you their message [1]

Spam in voice over IP (SPIT)

SPIT is the equivalent to junk mail in VoIP, for example we could receive unsolicited messages on our voicemail. It seems that even there is no SPIT on the wild now, experts expect this to be inevitable as soon as VoIP gets more popular

Blog Spam (web SPAM)

Blog o web logs are web applications that in some cases enable visitors to leave public comments. Blog Spam attacks the comment forms on the blog systems, filling the comments fields with links to Internet sites. They do this first to cheat the search engines and try to show their websites on the top of the page when you do a search for a specific topic in a search engine, and secondly to try to get someone who is reading the post, to click on the link. So every time a user clicks the link to go to the advertised web site, the link spammer makes some money. The "link spammers" or "search engine optimizers", as they like to be called, don't send all this traffic from their own PC because their ISP will find out easily and shut them out or the blog server would put them in a black list to block connection to their servers. That is why they use "Open Proxies" in Internet to send the messages they want but hiding their own connecting IP address.

There are several organizations like Google, Yahoo, MSN... who are trying to solve this problem, by adding a rel="nofollow" attribute to hyperlinks in comments, in this way webmasters and weblog owners can tell search engines that these links are effectively unreliable. So when the search engine sees the attribute (rel="nofollow") on hyperlinks, those links won't get any credit to rank websites in their search results. But it seems like the nofollow attribute will solve the problem of the search engines being cheated but not the problem of the web blogs being flooded with links.[2]

[1] See references 1, 2 and 5
[2] See references 4 and 5

Phishing

Phishing is a simple way to steal user credentials for authentication (i.e. password and credit cards details), through a specific online service, for example online banking. By pretending to be someone trustworthy who might naturally ask for this kind of information "Phishers" can induce users to give their confidential data over the Internet.
Normally this is done by organized fraudsters who use different techniques (such as Spam, Trojan horses, Spyware, Root-kits, Boots...) to commit fraud indiscriminately on a large scale.
Typically the user receives an email that looks like one his Bank might send him, except on this occasion he is being asked for his username and password for his Internet banking service.
In order to achieve this, fraudsters use different techniques to get the maximum amount of user credentials.
First they send massive amounts of unsolicited email normally using the same databases that spammers use, hoping to hit as many specific service users as possible. (for example a bank or an auction site).
Then they use social engineering to make sure the user assumes that the information is coming from the trustworthy entity. They do this by using the same domain name in the email and using the same look that the trustworthy entity uses.
Finally they ask for the credentials they are looking for, in a way that the user is forced to think that he is obliged to provide them, like saying that his account will be disabled if he doesn't provide this information immediately. They have very short time to receive the information before their fake servers are discovered and shutdown, so they use different techniques to force the user to send the information requested immediately.
An example of this kind of messages could be the following*: "If you don't respond within 24h after receiving this Mail Information your account will be deactivated and removed from our server (your account suspension will be made due to several discrepancies in your registration information as explained in Section 9 of the EBay User Agreement."*[3]
Also very often they explain to the user that the transaction will be completely secured using a secure connection, and also they will say that they would never ask us to send this sensitive personal information in an email because it could be dangerous, this is just to get victim trust before sending the personal information.

Normally they place a link in the email that will redirect the user to a fake web server. This web server does not belong to them, they are unsecured servers in Internet that have previously been hijacked to install the fake service and are configured to look exactly the same as the genuine site to an unaware user. Here is where the user will type all the credentials that the phisher is aiming for.

[3] See reference 6, from Anatomy of a Phishing Email

The credentials will be stored or transferred to the fraudsters for later use.

Phishers uses different methods to make sure that the information displayed to the unaware user looks convincing.
Some phishers may use a spoofed email address to make sure the victim thinks that the email originated from the reputable company.
They use the original images and text, but sometimes they just mirror the whole legitimate site with the same links and information, so the users think they are in the legitimate site.
Some phishing web sites use registered domains similar to the legitimate one like "http://www.popularbankone.com" instead of the genuine "http://www.popularbank.com" or using a subdomain like "http://www.popularbank.validate.com" where the real domain is "validate.com"
Then they normally use a collection form on a web site, this is used to collect all the information from the customer, using a format similar to the genuine service.
Sometimes they use incorrect URL or just IP for their websites, instead of bothering trying to disguise the domain name to make it more realistic. Also they can use obscure URLs by using the @ symbol and hexadecimal character codes to represent the numbers in the IP address, for instance http://popularbank@3468664375/obscure.htm is exactly the same as http://206.191.158.55/obscure.htm, this is because everything between "http://" and "@" is completely irrelevant. This is used only for authentication but if the site does not require authentication, the browser and server ignore this text

Other times they fake the address bar by opening a new browser window without address bar, and they build a new address bar using JavaScript commands with the URL that they want the user to see. Also they can use a text object with a white background over the URL in the address bar with just the URL they want to fake to cover the real URL being used.
They can also use pop-ups, in this case by clicking on the email link, the victim is sent to the fraudulent site, which firstly generates a pop-up window and then also redirects the main browser windows to the legitimate web site. So the victim sees the genuine website in the background, and makes him think that he is safe[4].

Methods of fighting email SPAM

Spam is growing every day at a steady rate; according to Symantec Spam Statistics dated January 2005 spam had increased 19% since the previous month. And it seems like the Product and the Financial category are the most popular kinds of SPAM worldwide. North America presents the highest amount of spam followed by Asia. We can also see that Spam makes up around 65% of all email messages worldwide.
Looking at these numbers it seems clear that ISPs, corporations and end users

[4] See references 7 and 8

need a way to reduce this amount of unwanted email that is clearly wasting company resources. This waste of recourses includes end user productivity, IT resources, email infrastructure, downtime due to virus or attacks, etc.
Fortunately most companies agree that they need to apply some kind of protection against SPAM, the solution should be a mix of solutions based on user education and technology implementation.

Educational approach

Users in a company should have a basic understanding of what SPAM is, why is it a problem, how to handle it and how to avoid it. For example, they should know not to use their corporate email addresses in Internet forums or when they register in unreliable websites, and if they are asked to give an email address to register to a new Internet service then they should also know to use a personal email address instead of the corporate email address. Also they should never try to unsubscribe from a spam distribution list, this is a trick the spammers use to confirm that this email address is in use. Equally of course, they should know not to buy anything advertised with spam as buying these products makes spam profitable for spammers.
Also, administrators of email servers need to know how to configure their SMTP servers to avoid to be used as relay servers and make sure they keep the systems and applications up to date with the latest patches.

Technical approach

It is clear that today there is no silver bullet anti-spam technology and protection has to be deployed using different technologies, each of these single technologies is good at stopping different kinds of spam.
There are a lot of products from different antispam companies that can help to solve the spam problem. But when choosing an antispam solution three things need to be taken into consideration, detection rate, false positives and the tuning/management needed to make the tool accurate. Although all three are very important, the false positive rate is the one that can ruin an antispam project, for example just missing one important email can make the company loose an important deal. That is why it is so important during the evaluation period to confirm that the false positive rate is reasonable for our company.

Bayesian filters: This is based on the idea that most events are dependent and the probability of the occurrence of one event can be inferred from the past occurrences of that event. To use this method we need to collect data from emails that are spam and valid emails, and give weight to each word depending of how often this word if found in spam and in valid emails
This word probability is calculated as follows: If the word "mortgage" occurs in

400 of 3,000 spam mails and in 5 out of 300 legitimate emails, for example, then its spam probability would be 0.8889 (that is, [400/3000] divided by [5/300 + 400/3000]) [5].

But in order for this to be effective it needs to be personalized for each company, otherwise the result will not be as effective.
After the database is created, when an email arrives it checks all the words and calculates the probability of the email being spam taking the whole message into consideration. So if it finds one word that could be spam and six words typical from valid email, the result would be "no-spam".
The Bayesian filter is constantly learning from new spam and new valid email, this is why if you implement a solution based on Bayesian filters you need to wait a couple of weeks before obtaining good results.

Reverse DNS Lookup: This does an nslookup query on incoming email source IP, if the domain name does not mach the one on the "from address" of the email, the message is rejected. This is a popular system but can generate problems with false positives because many DNS lookup entries are not or cannot be properly configured, like vanity domains.

Black/white list: This is the use of custom addresses that we either always want to allow (because we trust them) or we want to block (because they send us spam). The problem with black lists is that spammers can use a false "from" address, which makes handling this list very time consuming.

RBLs (Real-time Blackhole Lists): This checks all incoming email IPs against a list of IP addresses that are known for sending spam, and if it matches, it is identified as spam. There are different operators maintaining these DNSRBL lists, but it is important to choose the ones that best adapt to the organization, to avoid false positives. The problem with RBLs is that it only checks the incoming email IP, so it will block all the emails coming from that IP, even if it is valid email.

SPF Short for Sender Policy Framework, is an extension of SMTP that try to stop spammers from forging the "from" field. The SPF field has to be entered in the DNS systems of the participating domain with the IPs that are authorized to send emails from that domain, so when someone from this company send an email to someone else, this one can check if the IP of the sender is authorized by that domain.

Spam Databases: Some Antispam companies use honey pots to collect huge amounts of spam samples, and with all this information they use techniques such as hashing and fingerprinting to create a database that is updated to

[5] See references 9,10 and 11

customers´ antispam systems. Because spam messages are constantly changing, it means that updates have to be created very often to be effective.

Reputation services: This is a service that some Antispam companies provide and involves categorizing email senders by their originating IP address.
Inspecting through great amounts of messages they can identify machine IP addresses that are only sending spam. Similarly, they can also check systems that have been hacked and have mail server sending spam. After analysis they can make different lists, for example Open Proxy senders, Spam senders, Trojans sources, valid safe senders, etc, and can update the customers antispam systems.

Antivirus: Today a great deal of spam is generated by worms trying to spread themselves, so any antispam tool should have some kind of antivirus capabilities.

Traffic Shaping: This is a way of defending oneself against spam at the network layer. When these kinds of systems are implemented at the company gateway, all the senders receive Quality of Service identification according to the quality of their behavior. Spammers receive very poor quality of service, their message flow is restricted and the spam backs up on their own servers, on the other hand legitimate senders receive unlimited use of resources. Essentially, all sent messages eventually get through (although most spammers give up/time out) but spam is delivered at a drastically reduced rate.
It usually takes spammers more than a week to deliver a day's worth of spam; this is why it is the only way to effectively disrupt the spammer's economic model. Note that this system by itself does not block spam it simply slows it down, so it needs to be implemented in conjunction with another antispam technology.

Methods of fighting Phishing

Phishing is a problem based on the way authentications are made in the traditional online transactions. There are primarily two problems, firstly that the systems only use just one-way authentication, so only the user has to authenticate himself against the server, the server does not need to authenticate against the user, secondly the user has usernames and passwords that do not change constantly. So if someone manages to convince you that they are your trustworthy online bank and you give them your username and password, they can use these credentials as many times as they want until you notice the fraud is taking place.
Therefore a good solution would be to establish a two-way authentication, in such a way that before we authenticate our access to the Bank with our

credentials, the bank should authenticate our access with a password that is only known to the user and the bank. In this way the user can make sure that the site is what it says it is.
Another solution would be, to give a token to each customer so that every time they log into their online bank account, the password would be different, so even if the fraudster managed to convince you to put your credentials in his fake web site, this information would be useless by the next time he tries to use it. This solution sounds reasonable for Internet services like banking on-line. The only problem with this is that as the fraudster techniques evolve, they may use some kind of man-in-the-middle attack. This way the attacker would be listening to the communication between the user and the bank and when the unaware user authenticates himself with the token, the attacker could modify instructions, such as transferring money to another account or log the user out and keep the session for himself to use.
To avoid this from happening it is necessary to change from session-authentication to message-authentication: Some vendors offer a solution based on one-time password distribution using SMS messages, so when you order a transaction you receive a message via SMS with a description of the transaction and a unique code that you have to send back to confirm the order. In this situation it is really difficult for the attacker to compromise the SSL channel and the SMS channel at the same time[6].

Until an authentication solution is implemented, there are other ways in which financial institutions can minimize risks, they can educate users never to give their secret credentials, or they can offer real-time security alerts to make users aware of an incident, they can encourage them to report any suspicious activity, and they can offer them an online utility to scan their computers to look for virus, worm, Trojans. The possibility also exists to have a specialized company monitoring phishing messages in Internet and alerting when the financial institution is target of a phishing attack, this way they can start internal security procedures to avoid scam, in an early stage of the attack, they can take actions to close down the fake servers and temporally disconnect affected IPs. This monitoring also allows them to collect valuable information to prosecute the fraudsters. The idea behind this is to stop brand damage at the earliest stage if possible, in an attempt to minimize the number of users that receive the fake email. When the customer receives the email then the education of the users comes into play
These are easy best practices for end users to prevent online fraud:
- Do not click on hyperlinks within emails that are apparently from a legitimate company.
- Fraudulent emails use generic greeting like Dear Customer or Dear Member, or the email address as a salutation.
- Use Antivirus, antispam and personal firewall in the desktop
- Keep the OS and browser permanently updated
- Confirm that the web site where you are going to type your credentials has

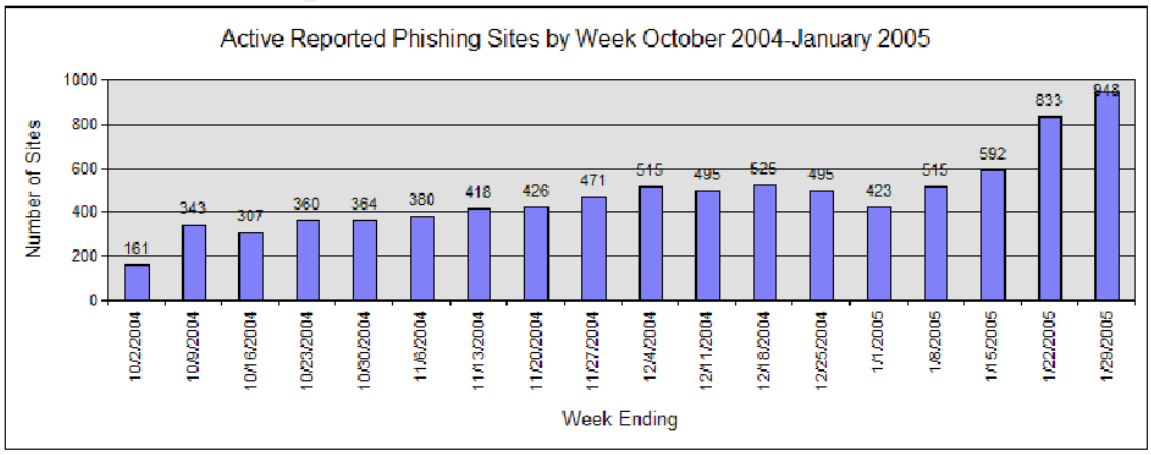
[6] See reference 12 and 13

"https" before the URL and have the security padlock icon on the bottom right hand corner

The Future of SPAM and phishing

We can expect to see email spam rising in the future, and even thought most countries are working on legal regulation to avoid spam, this is not going to help in the near future, so the best solutions at this moment are technology based.
Some experts expect the end of spam by 2006, but not everyone agrees with such an optimistic date. The end of email spam is expected by ISPs and enterprises using antispam solutions and the help of the law enforcement. But even if email spam problem comes to an end, it is reasonable to expect that spammers will then have to move on to other kind of spam like SPIM (spam for Instant Messaging), SPIT when VoIP gains more traction in the market and becomes more popular or other new kinds of spam using some new technologies.

On the other hand phishing doesn't appear to be coming to an end, on the contrary, we are seeing fraudsters using increasingly sophisticated methods. As we have seen it is very difficult to defend against phishing attacks, it is increasingly difficult for users to distinguish what is legitimate mail and what is not. Also we are seeing malicious code being used to get user keystrokes, and this malicious code does not come just via mail but also via IM or web servers.
Quite a large part of spam traffic is fraud email, according to Symantec's January 2005 Spam Statistics Report, 7% of spam traffic is associated with phishing[7].
If phising continues to rise at this rate [see graphic below from antiphishing.org], user confidence will soon be affected and this will have an impact on business and organizations that rely on transaction conducted over Internet.



[7] See reference 14

Graphic courtesy Tumbleweed Communications

Conclusion

We are experiencing a huge increase in Internet traffic due to these large amounts of unwanted emails being indiscriminately sent to mail servers. This has a direct impact on the resources ISPs and Corporations need to maintain good service levels, it also affects end user experience when they waste time cleaning up their inbox tray or when they are victims of fraud.
As long as spam and phishing proves profitable, there will always be someone wanting to make money from it, as a consequence ISPs, enterprises and end users will have to work towards minimizing its impact, by implementing any of the technologies here above described, not forgetting the important role of end user education on how to prevent fraud by knowing how to react against spam and phishing.

References

1. Symantec Response, Adware Buddylinks
http://securityresponse.symantec.com/avcenter/venc/data/adware.buddylinks.html

2. New Scientist. Spam being rapidly outpaced by 'spim'
http://www.newscientist.com/article.ns?id=dn4822

3. Marty Schultz, How to stop IM and SPIM abuse http://news.zdnet.com/2100-9595_22-5201055.html

4. The register. Interview with a link spammer
http://www.theregister.co.uk/2005/01/31/link_spamer_interview/

5. Google.Preventing comment spam http://www.google.com/googleblog/2005/01/preventing-comment-spam.html

6. Anatomy of a Phishing Email http://www.ceas.cc/papers-2004/114.pdf

7. How to Obscure Any URL http://www.pc-help.org/obscure.htm

8. FraudWatch International . Methods of Deception Used in Phishing Scams
http://www.fraudwatchinternational.com/internetfraud/phishing/website.htm

9. Symantec Brightmail AntiSpam
http://enterprisesecurity.symantec.com/products/products.cfm?ProductID=642%20

10. Gfi white paper http://www.gfi.com/whitepapers/why-bayesian-filtering.pdf

11. http://www.barracudanetworks.com/resources/docs/Spam_Techniques.pdf

12. Dávila, Jorge “En las Fuentes del phishing” SIC #63, pg 80 Feb 2005

13. The Future of Phishing by Dr. Jonathan Tuliani
http://www.cryptomathic.com/pdf/The%20Future%20of%20Phishing.pdf

14. Symantec Spam Statistics Report, January 2005

15. http://antiphishing.org/APWG_Phishing_Activity_Report-January2005.pdf