A report on the status of phishing on the web and sample of technology initiatives being used to stem the tide of phishing attacks

UNISA-TR-2006-03

Authors: Upasna Bechan
Supervisor: Alta van der Merwe
## Contents

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONTENTS</td>
<td>2</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>3</td>
</tr>
<tr>
<td>GRAPHICAL REPRESENTATION OF THE REPORT</td>
<td>4</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>5</td>
</tr>
<tr>
<td>WHAT IS PHISHING</td>
<td>7</td>
</tr>
<tr>
<td>WHERE DID IT ORIGINATE</td>
<td>7</td>
</tr>
<tr>
<td>WHAT FORMS OF PHISHING ARE THERE?</td>
<td>8</td>
</tr>
<tr>
<td>HOW DOES IT WORK?</td>
<td>11</td>
</tr>
<tr>
<td>HOW TO MODEL PHISHING ATTACKS?</td>
<td>13</td>
</tr>
<tr>
<td>WHY DO PEOPLE FALL FOR IT?</td>
<td>15</td>
</tr>
<tr>
<td>THE ANTI PHISHING MEASURES IMPLEMENTED BY YAHOO, EBAY AND EARTHLINK</td>
<td>17</td>
</tr>
<tr>
<td>SCANNING AND ALERTING SOFTWARE TO DETECT AND DEFEND AGAINST PHISHING ATTACKS</td>
<td>20</td>
</tr>
<tr>
<td>EMAIL VERIFICATION TOOLS</td>
<td>23</td>
</tr>
<tr>
<td>WEBSITE VERIFICATION SOLUTIONS</td>
<td>25</td>
</tr>
<tr>
<td>GUARANTEE OF LEGITIMACY OF MESSAGES AND THEIR SOURCE</td>
<td>28</td>
</tr>
<tr>
<td>TWO FACTOR AUTHENTICATION FOR CUSTOMER ACCESS TO FINANCIAL SERVICES</td>
<td>30</td>
</tr>
<tr>
<td>IDENTITY THEFT</td>
<td>34</td>
</tr>
<tr>
<td>CONCLUSION</td>
<td>37</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>39</td>
</tr>
</tbody>
</table>
List of Figures

FIGURE 1. GRAPHICAL REPRESENTATION OF THE REPORT LAYOUT ............................. 4
FIGURE 2. NEW PHISHING SITES BY MONTH OCTOBER-AUGUST ................................. 5
FIGURE 3. PASSWORD STEALING MALICIOUS CODE URLs ......................................... 6
FIGURE 4. PAYPAL PHISHING EMAIL ......................................................................... 9
FIGURE 5. THE PHISHING SITE LOADS UP AN ADDRESS BAR SPOOF AND A MIRROR OF THE LEGITIMATE PAYPAL.COM SECURE LOGIN PAGE ...................................... 10
FIGURE 6. DETAILED PHISHING ATTACK ..................................................................... 13
FIGURE 7. MODEL OF A PHISHING ATTACK .................................................................. 14
FIGURE 8. EXAMPLE OF A PHISHING ATTACK ............................................................... 15
FIGURE 9. PHISHING TOOLBAR: EBAY ACCOUNT GUARD ............................................ 17
FIGURE 10. THE EARTHLINK SCAMBLOCKER TOOLBAR ............................................. 18
FIGURE 11. EARTHLINK TOOLBAR SITE WARNING FEATURE .................................... 19
FIGURE 12. SPOOFGUARD ARCHITECTURE ................................................................. 22
FIGURE 13. SPOOFGUARD TOOLBAR ......................................................................... 23
FIGURE 14. STRONG TOKEN-BASED AUTHENTICATION ............................................... 31
FIGURE 15. MAN-IN-THE-MIDDLE ATTACK STRUCTURE ............................................ 33
FIGURE 16. IDENTITY THEFT LIFECYCLE .................................................................... 35
Graphical representation of the report

This report comprises 8 sections with Section 8 being a utility section for references. Sections 5 and 6 can be read in any order as their content is confined to their section boundaries.

Section 1: Introduction

Section 2: Background
- What is phishing
- Where did it originate

Section 3: Mechanics
- What forms are there
- How does it work
- How to model phishing attacks

Section 4: People
- Why do they fall for it

Section 5: Countermeasures
- The anti-phishing measures implemented by Yahoo, eBay and EarthLink
- Scanning and alerting software to detect and defend against phishing attacks
- Email verification tools
- Website verification Solutions
- Guarantee of legitimacy of messages and their source
- Two factor authentication for customer access to financial services

Section 6: Identity Theft

Section 7: Conclusion

Section 8: References

Figure 1, Graphical representation of the report layout
Introduction

This report begins with a description of what phishing is, how a phishing attack is conducted and why users fall for these attacks. It then goes on to explore a selection of technology initiatives being investigated in an attempt to handle the rising number of phishing attacks. The six areas of initiatives explored are:-

- The anti phishing measures implemented by Yahoo, eBay and EarthLink
- Scanning and alerting software to detect and defend against phishing attacks
- Email verification tools
- Website verification solutions
- Guarantee of legitimacy of messages and their source
- Two factor authentication for customer access to financial services

According to the Phishjing Activity Trends Report (August 2005), the number of unique phishing websites detected by the Anit-Phishing Working Group (APWG) was 5259 in August 2005, the highest number this far. This statistic helps to highlight the fact that phishing is a lucrative crime that will continue to grow unless effective counter measures are sought and implemented.

![New Phishing Sites by Month October-August](image)

**Figure 2. New Phishing Sites by Month October-August**
It is disturbing to note in the Phishing Activity Trends Report (November 2005) that the theft of passwords is increasing. This observation leads one to fear that the theft of passwords will lead to the theft of personal information that is accessible once the password is known.

Figure 3. Password Stealing Malicious Code URLs (APWG 2006, p. 6)

The illegal access and use of personal information is referred to as identity theft. Identity theft is discussed in Section 6 according to the following broad guidelines:-

- what kinds of identity theft there are
- what can it be used for.

The remainder of this paper is organized as follows. Section 2 begins with background information on phishing. Section 3 looks at the mechanics of how a phishing attacks is conducted. Section 4 attempts to understand why people are lured into a phishing attack. Section 5 explores a number of countermeasures and their effectiveness against the attack they are being used against. Section 6 is a brief introduction to problem of identity verification and includes a theoretical architecture that has been proposed in response to the identity theft problem. Section 7 is the conclusion and Section 8 contains a list of references.
What is phishing

Phishing is the term used by the online community to describe a form of social engineering whereby unsuspecting users divulge personal information (such as credit card numbers, usernames, passwords and social security numbers) to 3rd parties. The 3rd party then uses this information to commit credit card fraud and identity theft leading to financial losses by the owner of said personal information.

According to Rudd (2004), the term ‘phishing’ originates from the word ‘fishing’ where the fishermen use bait to attract fish like the criminals use email to attract online victims. The ‘f’ was replaced by ‘ph’ in recognition of ‘phreaking’, an early hacking method where the hacker would use someone else’s phone line for their own use.

Where did it originate

The earliest cited phishing attack is the one launched against AOL in the early nineties where criminals stole AOL passwords. Morrison and Nuttall (2004 :1) say “Phishing first came to light in 1996 when AOL subscribers began receiving bogus instant messages asking for their log-on passwords in order to update files. The aim was to avoid paying a subscription by using someone else’s account …”

Phishing was seen as largely a novice pastime but recently the activity seems to be more organised and launched against organisations for major financial benefit. Kerner (2004) quotes the Gartner survey of 5000 individuals which estimates the cost of damage to US banks and credit card issuers at 1.2 billion US dollars. There could be further financial implications as consumers lose confidence in transacting online.
What forms of phishing are there?

Identify theft can occur by a variety of techniques. The most widely known and most visible technique is sending the user an email which contains a link to a fraudulent website. Once the user is convinced to click on the link (See the section ‘Why do people fall for it?’ for a discussion on how this is accomplished), they are redirected to the spoofed site where they enter their personal details. The details are then sent to the criminal’s web server/database.

Below is an example of an email from APWG (2005) that was used in a PayPal phishing attack.
Dear PayPal,

We recently noticed one or more attempts to log in to your PayPal account from a foreign IP address.

If you recently accessed your account while traveling, the unusual log in attempts may have been initiated by you. However, if you did not initiate the log ins, please visit PayPal as soon as possible to verify your identity:


Verify your identity is a security measure that will ensure that you are the only person with access to the account.

Thanks for your patience as we work together to protect your account.

Sincerely,
PayPal

-----------------------------------

PROTECT YOUR PASSWORD

NEVER give your password to anyone and ONLY log in at https://www.paypal.com/. Protect yourself against fraudulent websites by opening a new web browser (e.g. Internet Explorer or Netscape) and typing in the PayPal URL every time you log in to your account.

-----------------------------------

Please do not reply to this e-mail. Mail sent to this address cannot be answered. For assistance, log in to your PayPal account and choose the "Help" link in the header of any page.

PayPal Email ID PP321 VIYJBXRJ0KBEFXKFPNXZDBBYDMZK WBGBGYYHOL

Figure 4. PayPal phishing email

The link referred to in the email above takes the user to the spoofed website listed below:-
Figure 5. The phishing site loads up an address bar spoof and a mirror of the legitimate paypal.com secure login page
A variation on the email scam described above is to embed an html form in the email that is sent out. The user then enters their credentials in the email form which has been embellished with logos/colour schemes to accomplish the appropriate branding.

Another effective technique to illegally obtain user information is to send the user a worm that installs key logger software on the victim’s personal computer (PC). All input data is then sent to the criminal’s website where the required information is filtered out. This technique is more difficult to detect hence the user is unaware that anything is wrong until fraudulent credit card transactions, etc. occur.

A further technique is to create a pop up window over the legitimate website. The user then enters their login credentials on the pop up thereby giving the criminals the information they need. The user does not realise what has happened as they still have the legitimate website in front of them.

A recent trend in phishing attacks is targeting Internet Relay Chat systems (IRC’s). During a chat session the user downloads a file with attachments that install key logging spy ware on his personal computer (PC).

**How does it work?**

Below is a detailed analysis of what happens during a phishing attack from Tumbleweed Whitepaper (2004 :3,4):-

Assume the following actors and components of the attack:

- The phisher (P)
- The recipient (R)
- The company whose domain is being spoofed by the phisher (C)
- The phisher’s email server (Ps)
- The recipient’s email server (Rs)
- The recipient’s email client (Rc)
1. P generates a fraudulent email with content that looks just like a legitimate email from C to all its customers. The colors, graphics, text treatment, and composition are identical to what C uses to normally contact its customers. There is no way for C to technically prevent P from creating this type of content. The message in this email is particularly insidious as it describes how “A recent set of phishing attacks have corrupted our customer account database at C. Please help us reinstate your account at a secure website provided by the C Security Service by clicking on the link below”. P then inserts the email address customerservice@C.com in the FROM field of the email.

2. P sends this email to as many email addresses as he can get a hold of using server Ps. He may have done previous spam attacks to understand which email addresses are likely C’s customers to reduce the number of phishes he must send to be effective. The domain of Ps happens to be Csecurityservice.com. The address P uses in the MAIL FROM address of the email is customerservice@Csecurityservice.com. This is so that any bounces from invalid recipient email addresses are sent to P’s servers and not C’s. This will prevent C from being Joe Jobbed and noticing a large number of bounces coming from email it knows it did not send.

3. R happens to be a customer of C and has heard about these phishing attacks on the Internet. R’s email server (Rs) receives a connection from Ps, accepts it, and stores the message in R’s email inbox. R uses his email client (Rc) to download the message. When he sees P’s email, he inspects it closely. The address that Rc displays, customerservice@C.com, has C.com in it, and the “customerservice” string to the left of the “@” matches another email he got from C a few months back regarding a product return he made. R trusts this address as legitimately coming from C.

4. While R thinks he’s savvy to phishing attacks, the content of P’s email takes him off-guard. “Why sure”, R thinks to himself, “These phishing attacks probably are wreaking havoc on corporate account databases. I read about it in the papers. I better make sure my account is not ruined, because I use C’s site often.” R clicks on the link and is taken to a Web site.

5. R is taken to https://www.Csecurityservice.com/accountreinstatement. R remembers that the email said the site to reinstate his account was provided by a “C Security Service”, so the syntax of this URL makes perfect sense to him. The SSL connection his browser tells him has been made further convinces him that the site is legitimate. The form on the site asks for just the right kind of personal information that C would ask of its customers in order to reinstate an account. R fills out the form and hits the submit button. The phisher has now succeeded in stealing personal information from R.

6. R realizes some time later that he’s been phished. C may discover P’s attack from R or other victims’ reports. By the time C is able to get P’s ISP to shut down the https://www.Csecurityservice.com/accountreinstatement
URL, P has already gathered hundreds if not thousands of sets of passwords and other personal account information. The fact that P can never again use https://www.Csecurityservice.com/accountreinstatement as a Web site or even customerservice@Csecurityservice.com as the MAIL FROM address in another phishing attack (due to updated blacklists in anti-spam servers) is beside the point. He’s made a killing with stolen identities he already has. P can easily move on to the next company’s set of customers to launch a phishing attack. He might even try phishing C’s customers again with an entirely different set of legitimate-looking domain names and email content.

Figure 6. Detailed phishing attack

How to model phishing attacks?

The purpose of using a model as opposed to descriptions (See the section How does it work?) is the model is able to capture “a variety of attacks in a uniform and compact manner”, Jakobsson (2005 : 4).

The graphical model allows the analyst to visually quantify the threat to the system under review. The phishing attack is modelled by a phishing graph “… in which nodes correspond to knowledge or access rights, and (directed) edges correspond to means of obtaining information or access rights from already possessed information or access rights … ”, Jakobsson (2005 : 2).

Below is an annotated example of a phishing graph with detailed explanations of the labels below.
Vertices can correspond to access to either information or a resource; the distinction in the graph is implicit by the associated description. Actions are represented by edges, so two vertices are connected by an edge if there is an action that allows a user with access to one vertex to obtain access to the linked vertex. Two vertices may be connected by multiple edges, corresponding to different actions allowing the transition between them; this is referred to as a disjunction in Figure 4 above. A conjunction of actions occurs when a number of edges can be merged into one and ends up in one vertex.

Edges are labelled with descriptions of the effort, probability and circumstances under which the transition will succeed. The target refers to the vertex the attacker is attempting to gain access to.

Figure 5 is an example of a phishing attack described in the graphical notation of Jakobsson. The context of the example is as follows:- Access to a newly opened bank account can be obtained by submitting the date and amount of the last deposit as opposed to a password. The date, amount and account number is assumed to be known for this example.
Figure 8. Example of a phishing attack

Access to the account is represented by vertex v1. Knowledge of the victim’s salary is represented by vertex v2, and the edge e251 corresponds to guessing the level of withholding and percentage of 401(k) contributions. Knowledge of the victim’s marital status corresponds to vertex v3, and the edge e351 is the probability of the tax refund check being deposited alone. Vertex v4 corresponds to access to performing a payment to the victim, and the edge e451 corresponds to the action of performing the payment. Vertex v5 corresponds to knowledge of the account number.

Why do people fall for it?

According to Tumbleweed (2004:1) “… the phisher is counting on the fear, guilt or general willingness of unsuspecting victims to trust the e-mail’s contents and follow its instructions”.

Ironically it's usually the fear of phishing attacks and fraudulent activity that is cited as the reason for needing the user information to be validated. The user is also told that it’s a matter of urgency that the update is carried out. This does not give the user time to validate the email or its contents and the
criminal gets the information they need before the website is detected and shut down.

According to Drake, et. al (2004), users trust the email they receive for the following reasons:-

a) the company being mimicked is a reputable company e.g. Paypal
b) the spoofers emulate the company’s visible branding
c) the email or spoofed website may contain links to sections of the legitimate website
d) the from address appears to be from the legitimate website

People also trust that the certificates authority (CA) has carried out the due diligence. This may not always be true. According to Hall (2005), there are 3 areas of weakness in this process:-

a) Manual vetting process would not be able to spot fake documents hence making it easy for fraudulent companies to obtain certificates.
b) The vetting process was not standardised across states and countries hence making it more error prone.
c) Some CA’s outsource the identity verification process to dubious subcontractors who vouch for the business without doing the proper checks

So when the user looks at a web site and sees the padlock on the toolbar at the bottom of the browser, this is no guarantee of the business’s validity. What the user has to do is double click the padlock icon, navigate to the ‘Subject’ section and verify the data fields displayed from the digital certificate. This is too complicated and more often than not the data therein is not accurate.

The solution to the error prone manual vetting process is Second Generation Automated Vetting. The reasons cited for this process working are:-

a) domain (web address) is confirmed by the applicant in real time
b) real time email validation
c) real time telephone validation
d) sophisticated fraud-detection algorithms
The anti phishing measures implemented by Yahoo, eBay and EarthLink

Yahoo proposed the DomainKeys infrastructure that verifies the domain of an email sender. This verification helps in the war against phishing by helping to curb the proliferation of spam. Shor (2005) describes DomainKeys Identified Mail (DKIM) by “Each message gets a signature, which can then be checked for authenticity by verifying the originating domain and whether the message has been tampered with.” The process of using DKIM is described by Germain (2005) as follows. The ISP or e-mail gateway authenticates the message sender. Then the message must pass a reputation score. Thereafter the Domain Name Server (DNS) is used to verify that the encrypted email address signature came from the stated sender. Shor (2005) describes DKIM as “…help to mitigate address spoofing…but does not eliminate either (phishing or spamming) entirely”.

The eBay toolbar is a browser plugin whose primary role is to keep track of auction sites. However the toolbar also helps prevent fraud in the following manner according to Dhamija and Tygar. “AccountGuard”, a feature of the eBay toolbar monitors the domain names that the user visits and provides a warning by changing the colour of the tab on the toolbar. The toolbar, usually grey, turns green if the user is on an eBay or PayPal site or turns red if on a spoofed website. This toolbars effectiveness is only as strong as the list of spoofed websites known to eBay and PayPal.

Figure 9. Phishing Toolbar: eBay Account Guard
EarthLink has made available a free anti-phishing application called ScamBlocker. ScamBlocker is a feature of the EarthLink toolbar. The purpose of ScamBlocker is to prevent users from accessing phishing sites. According to Tynan (2004) ScamBlocker attempts to stem the phishing tide in the following manner. ScamBlocker automatically downloads a list of known phisher sites, sourced from its own list, EBay (online auctioneer) and Brightmail (antispam vendor). When the user tries to access a fraudulent site, ScamBlocker redirects them to a page supplied by EarthLink. The EarthLink page then gives the user the choice of either proceeding to the scam site or reporting the incident to the EarthLink abuse team. The abuse team will then attempt to get the site’s host to shut down the site.

The benefit of using ScamBlocker is the user is alerted to the fact that they are about to visit a known phishing site. Unfortunately ScamBlocker is not foolproof as the feature is only as effective as the list it has access to.

The following list of figures from Lininger and Dean (2005) illustrate now the EarthLink ScamBlocker toolbar works:-

Figure 10. The EarthLink ScamBlocker toolbar
Digital Impact (2005) defines an effective anti-phishing solution as one that meets the following three requirements:

a) fast and accurate detection system  
b) rapid alert system to inform affected parties  
c) effective response mechanism

The solution being marketed by Digital Impact comprises three tools that meet each of the above requirements.

**Fraud Detection-Alert-Response** system (Fraud-DAR) comprises of Phishing Site Detection, Email Phish Detection, ISP Alert and Response, Customer Alert System and Real Time Reporting elements. The system allows the user to detect a possible attack early and inform potential victims.

**Personal Authentication System** is made up of the Recipient-Selected Authenticity Phrase and Sender-Selected Authenticity Phrase elements. This system allows the user to distinguish legitimate mail from fraudulent email because each email can include a unique, personal authentication item, for example an image, code, phrase, rewards balance, etc..

**Email Verification Database** comprises the Customer Service Email Database which allows the user to query the database using the customer’s email address and email code identifier and quickly determine whether an email was sent by the company or is fraudulent.
Scanning and alerting software to detect and defend against phishing attacks

All scanning and alerting mechanism look for patterns of previous attacks and attempt to stop any future attacks. This section explores intrusion detection systems for email filtering and web browser support.

Email auditing is the process of checking the email after the actual transmission. Email filtering is the process of intercepting and checking email during transmission.

Email auditing can be achieved by using intrusion detection systems. An intrusion detection system automates the monitoring and analysis process. As this is a reactive, process, the monitoring and analysis will occur after the email is sent.

An intrusion detection system (IDS) is a device that monitors activity to identify malicious or suspicious events. Intrusion detection systems are either signature or heuristic based. Signature based systems perform simple pattern matching and report situations that match a pattern corresponding to a known attack type. Heuristic systems build a model of acceptable behaviour and flag exceptions to that model.

Intrusion detection devices can either be host or network based. Host based systems operate on a host to detect malicious activity on that host. Network based systems operate on network data flows.

Intrusion detection software builds patterns of normal system usage triggering an alarm any time the usage seems abnormal.

An active intrusion detection system may be used to get the signatures of known phishers and then the IP can be blocked. IP’s can also be obtained by the use of a honeypot.
“Honeypots simulate one or more network services, hoping that an attacker will attempt an intrusion. A honeypot is configured to interact with potential hackers in such a way as to capture the details of their attacks. A properly configured honeypot monitors traffic passively, doesn’t advertise its presence, and provides a preserved prosecution rail for law enforcement agencies.“ (Lininger and Dean 2005).

Email auditing can also be achieved by scanning audit logs of email activity. Logging is a reactive process which can occur at the application, host or network level, meaning that an email trail can be picked up at any point in the communication chain. Most email systems use access control and /or passwords to authenticate the user and determine what activity the user can perform. The use of access control and passwords can be audited at the application, host and network level hence the users activity can be monitored.

Email filtering can be achieved by the use of firewalls where specific data is either blocked or allowed to go through. The firewall can either be implemented at the host or network level. At the host level filtering occur for one server whereas at the network level filtering occurs for all servers behind that firewall. Personal firewalls can also be used for filtering on a per user basis.

Phishing, also known as web spoofing is considered by Chou et. al. as a special case of intrusion detection and they propose a browser plugin (called SpoofGuard) to help combat web spoofing.

SpoofGuard is an Internet Explorer (IE) plugin that accessed the IE history file and three other files stored in the user profile directory:-

1. file of host names of email sites
2. file of hashed password history
3. file of hashed image history
Below is a diagrammatic representation of the SpoofGuard architecture from Chou et. al., followed by a summary of the characteristics of the SpoofGuard architecture:-

Figure 12. SpoofGuard architecture

- SpoofGuard uses domain name, url, link and image checks to evaluate whether a page is being used in a spoof attack
- SpoofGuard uses history logs to make intelligent decisions about user behaviour
- SpoofGuard intercepts and evaluates user posts
- SpoofGuard calculates and evaluates the spoof index of a page
- SpoofGuard compares post data to previously entered passwords from different domains

When the spoof index is above a user-specified threshold a pop up window is displayed with an additional warning. The user then has the choice to stop what they are doing for continue sending data to the third party site. If the user does not like pop ups a less intrusive tool bar is also available to give the user information about the page.
The Options button allows the user to configure the tool. The traffic light (red, yellow, green) gives an indication about the current page. If the user clicks the traffic light then additional information about the page is displayed.

Chou et. al. evaluated SpoofGuard against similar characteristics as intrusion detection systems and found the results very favourable.

**Email verification tools**

According to the Anti-Phishing Working Group (APWG) as quoted in Lininger and Dean (2005), a two step email authentication standard could stop 85% of phishing attacks in their current form.

Lininger and Dean (2005) list the following four as the main contenders for authentication:-

- **SPF (Sender Policy Framework)**
  Checks the ‘envelope sender’ of an email message-the domain name of the initiating SMTP server
- **Sender-ID**
  Checks after the message data is transmitted and examines several sender-related fields in the header of an email message to identify the purported responsible address
- **DomainKeys**
  Checks a header containing a digital signature of the message. It verifies the domain of each email sender as well as the integrity of the message.
- **IIM (Cisco Identified Internet Mail)**
  Adds two headers to the RFC 2822 message format to confirm the authenticity of the senders address.

Sender Policy Framework (SPF) was formerly known as Sender Permitted Form and is an extension of the Simple Mail Transfer Protocol (SMTP).
According to Lininger and Dean (2005) “When a user sends you mail, an email servers connects to your email server. When the message comes in, your mail servers can, based on SPF published addresses of its email servers, tell if the server on the other end of the connection actually belongs to the sender.”

Sender-ID, proposed by Microsoft, is the culmination of two previous protocols, namely **Caller ID** and **Sender Policy Framework** (SPF). The aim of the Sender-ID protocol is to only allow authenticated messages to reach the receiver. According the Germain (2005) Sender-ID attempts to achieve its aim in the following manner. The sender sends an email message to the receivers inbound mail server. The receiver’s server checks for a record of the sending domain published in the DNS record. The inbound email server determines if the sending email server’s IP address matches the IP address published in the DNS record.

DomainKeys is used to verify the domain of the email sender and the integrity of the message. DomainKeys uses public key encryption technology to achieve the verification. The process of how DomainKeys are used can be described in two parts, Sending DomainKey email and Receiving DomainKey email according to Lininger and Dean (2005).

Sending DomainKey email involves performing a secure hash of the email contents using the SHA-1 algorithm; encrypting the result using a private key with RSA algorithm; and encoding the encrypted data using Base 64.

Receiving DomainKey email involves the server using the name of the domain from which the email originated to perform a DNS lookup to get that domain’s public key. The receiver decrypts the hash value in the header and recalculates the hash of the body. If the two values match then the receiver can be confident of the origin of the email.

When comparing Sender-ID and DomainKeys Gillis as quoted in (Germain 2005) commented that “Sender-ID is very lightweight but is already very widespread. Domain Keys is definitely stronger because of the encrypted signature.”
Cisco Identified Internet Mail (IIM) was designed by Cisco System to help identify fraudulent email. IIM is a signature based email authentication standard that adds two headers to the message format: IIM-Signature and IIM-Verification. Lininger and Dean (2005) state that “To establish the authenticity of an email message, IIM verifies that the message sender is authorized to send messages using a given email address and that the original message was not altered in any consequential manner.”

**Website verification Solutions**

Fraudsters prey upon the trusted brand relationship well known companies have with their customers; Website Verification Solutions try to answer the question of how to distinguish authentic web sites from copycats.

**Client Side Solutions**

Ollman (2004) describes the following three hidden attacks that are used to manipulate the display on the victims web browser:- hidden frames; overriding page content and graphical substitution. One small step towards countering these threats is to be aware of the browser specific visual clues of graphical substitution that Ollman (2004) lists, viz. “…the URL presented within the browsers URL field, the secure padlock representing an HTTPS encrypted connection, and the zone of the page source.”

The fraudulent website has to either be a full copy of the legitimate site or make reference to the real site for part of the graphics/content. Developers can make the fraudster’s task much more difficult by employing Image Cycling as suggested by Ollman (2004). Image Cycling is the technique of uniquely naming and recycling images periodically. Any fraudulent site that references an image by the old name could have content supplied that displays a warning to the user. A variation of this technique is to extend the name to include the user session.
Another graphical security technique is to use an image hash rather than the actual image on the legitimate site, Chou et. al. (2005)

Ollman (2004) also describes the use of “…agent-based bots to monitor URL’s and web content from remote sites, actively searching for all instances of an organisations logo, trademark or unique web content.” Obviously once the fraudulent website content is detected, the owner can take remedial/legal action.

The Comodo Website describes Content Verification Certificates (CVC’s) as an effective means of verifying website content. The website content is first checked to ensure that it’s verifiable (see requirements below) and then an X509 compliant certificate is issued. There is a high degree of confidence in the CVC’s as they are built on the public key infrastructure. The following are requirements for web site content to be verifiable according to the Comodo website:-

- Suitably complex such that it cannot easily be spoofed (No cut and paste possibilities)
- Directly linked (bound) to the web page (URL and or IP) upon which it is to be displayed
- Given a validity period related to its usage.

Chou et. al. (2005) presented three groups of client side tests that can be used to distinguish spoof pages from legitimate pages, i.e. stateless, stateful and methods that evaluate outgoing html post data:-

Stateless methods are be used to evaluate the current page to determine if it is suspicious or not. Examples of stateful methods would be: - The URL of the page is checked for correct form and any links on that page are also checked. Any page that requests a password can be checked for https and whether the site is using a valid certificate.

Stateful methods are used to evaluate a page based on previous user activity. The domain of a page is checked for a match against previously visited pages
or a close match in the domain name. Chou et al. (2005) achieves this check by calculating the Hamming (edit) distance. The referrer page can also be used as an indicator. As most phishing attacks are initiated by email, if the referrer is an email site then this raises the suspicion level. Methods that evaluate outgoing html post data includes hashing and storing sensitive data in a database so that any outgoing post data is checked against the password hash to detect password leakage.

**Server Side Solutions**

Chou et al. (2005) describes the following server side methods that can be added to the arsenal for website verification:

- Marking form fields with confidentiality tags will allow security tools to track the correct information thereby reducing false alarm rate. The confidentiality tag can also be used by browsers to flush cookies if the user had visited a site that required confidential data input.

- Image tagging is implemented by adding a tag to the page to mark the image as not to be used on pages other than pages belonging to the domain. If the tag is found on pages outside the domain then the page is flagged as suspicious.

- The use of password hashing and site specific salt will produce distinct passwords for distinct sites. The salt has to be a unique value and Chou et al. (2005) recommends using the site domain name as the salt. Since users tend to use the same password across sites, once a phisher has cracked one site they try the same password on other sites so password hashing and the site specific salt is an effective countermeasure.

Another techniques that Chou et al. (2005) describe but are not in favour of are collaborative methods that rely on the user informing a central server of a spoofed site, the central server then alerts all plug-ins to block this page.

**HIPs**

The implementation of the Human Interactive Proofs (HIPs) criteria proposed by Rachna and Tygar (2005) is called Dynamic Security Skins (DSS) and provides a solution that allows the “remote server to prove it’s identity in a way
that is easy for a human user to verify and hard for an attacker to spoof.”
(Rachna and Tygar, 2005:p136)
According to Rachna and Tygar (2005) “… (HIPs) allow a computer to
distinguish a specific class of humans over a network. “. They extend the
definition of HIPs to allow a user to issue a challenge to the computer. The
characteristics of the challenge are:-
- “be easy for a particular class of computer to pass
- be hard for other computers to pass, even after observing a number of
  successful authentications
- produce results that are easy for a human to verify
- use a protocol this is publicly available
- not require the user to have specialized tools”

**Guarantee of legitimacy of messages and their source**

To guarantee the legitimacy of messages once must look at the question of
determining whether the website content is from the provider who the site
purports is the provider. Fraudster use many tactics to spoof a site, including
- official looking and sounding email
- copies of legitimate corporate emails with minor url changes
- use of fake “Mail From:” addresses and open mail relays for disguising the
  source fo the email

The popularity of Instant Relay Chat (IRC) and Instant Messaging (IM) has led
to these applications being enriched with dynamic embedded content, for
example URL’s and graphics. Many of the phishing tricks used on web sites
can now we used on IRC and IM applications as well.

**Validating Official Communication**

Official communication sent be an organisation can be constructed in such a
manner as to give the recipient visual clues to its authenticity or lack thereof.
Some methods of helping the recipient judge the authenticity of a message,
identified by Ollman (2004), are email personalisation; visual/audio
personalisation of messages; previous message referral and digital signatures.

Email personalisation
The organisation personalises any email with information that is only shared between the organisation and its customer, hence generic email will immediately cause the customer to be suspicious. However the organisation has to vigilant about who can access to the shared customer information. It’s possible to use visual and audio data for the purposes of personalisation.

Previous Message Referral
The organisation can, in current correspondence reference the last correspondence with its customer. Ollman (2004) identifies the following suggestions:-
- Clearly referencing the subject and date of the previous email
- Sequentially numbering the emails

Digitally Signed Email
By digitally signing the message the recipient can be confident of the originator, i.e. the organisation/person that signed the email, leading to a better trust relationship. However the onus is on the recipient to verify the signature thereby ascertaining the true source of the message. Ollman (2004) issues a word of caution though, there is nothing stopping the phisher from creating a public/private key pair and digitally signing his mail.

Mail Server Authentication
There are 2 ways to ensure you only communicate with message senders that are trusted according to Ollman (2004). Either you verify the IP or range by executing a reverse resolution of Domain information or you use secure SMTP. The first method will cause emails to be dropped if the sending domain cannot be verified while the second method will fail at the establishment of a secure connection. The purpose
One of the best practices recommended by ASTA (2004) to ensure that the source of a message is known is to reconfigure any mail servers that are configured as open relays to be secure relays. Open relays allow messages to pass through the mail server obscuring the sender from the receiver.

Another recommendation from the Anti-Spam Technical Alliance (2004) is for ISP’s are to ensure anyone wanting to send email must have a valid account on the system. The ISP’s can also impose rate limitations on outbound mail traffic to curb spam that is being used as a device to spread viruses/worms. Anti-Spam Technical Alliance (2004) suggests “The limits should be based on To/Cc/Bcc recipient counts per unit of time from end user account or server IP address.”

Domain spoofing obscures the origin and true sender of email by forging the sender address and domain name. It is widely acknowledged that the IP address is the only trustworthy element of the email header. Anti-Spam Technical Alliance (2004) suggests one approach to counter domain spoofing is to validate the domain information in the header against the IP of the domain. The IP is them matched against a publicly known list of IP that are allowed to send email on behalf of that domain, obviously any deviation will create suspicion of a security risk.

Anti-Spam Technical Alliance (2004) also suggested Content Signing (CS) technology will be valuable in verifying the sender’s identity and the message contents. CS makes use of public key/private key pairs to digitally sign messages and achieve the authentication.

Two factor authentication for customer access to financial services

The case for two factor authentication is strongly stated by Ollman (2004) as “… to clearly determine consumer identities so online businesses can avoid the costs of being defrauded and dramatically reduce the overhead costs of today’s manual methods of fraud and theft prevention.”
Two-factor authentication technology refers to the dual authentication mechanism of *something you know* (a password) and *something you possess* (a *token*). The something you possess could be a physical device, (like a smartcard or key-fob) or a single-use or time-dependant password.

The purpose of two factor authentication, according to Ollman (2004) is to “…create strong (one-time) passwords that cannot be repeatedly used to gain entry to an application.”

![Figure 14. Strong token-based authentication (Ollman 2004, p.34)](image)

Hardware tokens devices currently available are of the Challenge-Response method or the SecureID devices from RSA security described below.

Kumar (year unknown) describes client certificates as being able to provide strong authentication in a web application. A Smart Card implementation of client certificates provides “a secure and mobile platform for authentication”, Kumar (year unknown).

**Example: RSA SecureID Consumer Protection Solution**
RSA Security Inc white paper (2004) describes the RSA SecureID Consumer Protection Solution implementation of two factor authentications in the following manner:-
- A hardware device, the RSA SecurID authenticator is issued to each person that needs secure access
- The RSA SecurID authenticator generates a new, unpredictable code every 60 seconds.
- Two factor authentication is achieved by combining the authenticator with a PIN
- Persons needing secure access then combine something they know, i.e. the PIN with something they possess, i.e. the constantly changing code on the authenticator

The benefits of two factor authentication in the context of the RSA SecureID Consumer Protection Solution are as follows:-
- The one time password changes every 60 seconds so a phisher cannot steal and use old codes
- There is no requirement to install additional software
- The authenticator is mobile hence secure access can be achieved from any device connected to the internet
- Users don’t have to remember multiple passwords
- Businesses can better protect themselves against fraud hence re-establishing trust in e-commerce

Schneier (2005) argues that two factor authentication is no longer adequate protection when transacting over the internet as “the real threat is fraud due to impersonation…Two factor authentication will force criminals to modify their attacks, that’s all”. Criminals have modified their approach by using Man-in-the-middle and Trojan attacks to render two factor authentications useless.

Man in the middle attacks
The attacker places an intermediary between the victim’s browser and the real web server to take control of the victim’s resources and confidential data and proxies all communication. Ollman (2004) lists the following 4 attack vectors for Man-in-the-middle attacks:- Transparent Proxies; DNS Cache Poisoning; URL Obfuscation and Browser Proxy Configuration
The man-in-the-middle attacks are successful because the attacker is proxying all communication between the users and secure resource so the attacker can easily insert their own transactions along with the user initiated transactions.

Trojan attacks are successful because the attacker uses the Trojan to piggy backs on the user’s legitimate session with the secure resource.

The use of strong two factor authentication opens up opportunities for creating communities of user and sites that trust each other or as RSA Security Inc white paper (2004) describes them “federated consumer identity protection”. The benefits of creating such trusted domains are:-

- shared cost of implementing security across multiple systems
- ease of and decreased time to market for a new site as they can now leverage of a proved authentication system
- consumers have one way of authenticating themselves across multiple systems

Two phase login

Ollman (2004) describes a two phase login process that can make the authentication process more secure. Phase one involves the input of data that may be common knowledge, e.g. the account number and login name. Once the user gets past phase one, they are required to input 2 or more unique pieces of authentication information. Phase two can also present a personalised graphic that potentially acts as a watermark on that page.
Identity Theft

Van Oorschot and Stubbebine (2004) define identity theft as “… the unauthorised use and exploitation of another individual’s identity-corroborating information (e.g. name, home address, phone number, social security number, bank account numbers, etc.)” The purpose of this type of crimes is to obtain enough information to either apply for new official documentation in the victim’s name or use the existence of the identity of the victim to conduct illegal activities.

The Identity Theft Resource Centre (2005) cites Gartner Research and Harris Interactive who estimate approximately 7 million people falling victim to identity theft in the prior 12 months.

They also cite a GAO study on the costs of identity theft in America “A GAO study on identity theft (GAO-02-363, issued March 2002) discussed costs to federal agencies -- The executive office for U.S. Attorneys estimated cost of prosecuting a white-collar crime case was $11,443. The Secret Service estimates the average cost per financial crime investigation is $15,000. The FBI estimates the average cost per financial crime investigation is $20,000”

These 2 statistics reinforces the severity of the problem and the seriousness with which counter measures need to be sought.

Identity theft is becoming increasingly popular because it is a relatively easy crime to commit and the rewards (monetary) are substantial. Van Oorschot and Stubbebine (2004) believe that the reason personal and financial information is so easily stolen is it’s very easy for the fraudster to duplicate personal information but very difficult for the victim to know that a duplicate has been made. When the fraudster applies for duplicate documentation using stolen personal information there is no mechanism for the legitimate owner to know that this process is taking place.

The Special Interest Group (SIG) on Identity Theft in the Liberty Alliance Whitepaper: Identity Theft Primer (2005) have categorised identity theft into the following three groups:-
‘True’ name identity theft - the fraudsters use stolen personal information to obtain new accounts and services

Account takeover – the fraudsters use the stolen information to gain access to the victims legitimate accounts and services

Criminal identity theft – the fraudsters use the stolen personal information to evade having criminal record in their own name

For Identity theft to be viable the information must first be accessed and then used for nefarious purposes. The SIG on Identity Theft believe that identity theft is not an ad hoc crime of convenience, and can be effectively used in a planned attack using the steps in the six phases described below

Figure 16. Identity Theft Lifecycle (Liberty Alliance Project 2005, p.7)

Phishing, key loggers and screen grabbers have proven to be effective means for the fraudster to obtain this “identity-corroborating information”. Users readily disclose this information on a spoofed site as there is inherent trust in the requesting authority, i.e. the bank whose logo/banners are prominently displayed.
Van Oorschot and Stubbebin (2004) believe key logging now rivals phishing as the tool of choice to illegally obtain personal information. Lemos (2004) as cited in Van Oorschot and Stubbebine (2004) describe the example of the Bankhook.A Trojan that recorded sensitive information prior to SSL encryption and mailed the data to a remote computer.

The Liberty Alliance Project (2005) advocates the three pronged approach of technical, operational and policy countermeasures to identity theft. Two technical approaches are described below, the use of activity monitoring and the theoretical architecture proposed by Van Oorschot and Stubbebine (2004). The operational approach necessitates the agreement to and widespread implementation of contracts and best practices. The policy countermeasure requires the legal and regulatory compliance of institutions that store and process personal information.

Many systems and companies attempt to recognize identity theft by monitoring activity/transactions made by an individual. The occurrence of transactions that fall outside the normal pattern of behaviour is then flagged as suspicious and warrants closer inspection. This approach is effective if the normal pattern of transacting is well defined and we are dealing with a transaction based system like credit card activity.

The solution proposed by Van Oorschot and Stubbebine (2004) is an architecture that combines “… a physical location cross-check, a method for assuring uniqueness of location claims, and a centralized verification process.” The following is an overview of the proposed system:-

- Every user has some device that is used to securely determine their location; Van Oorschot and Stubbebine (2004) refer to this as the “heartbeat locator”. Examples of heartbeat locators would be cell phone and wireless personal digital assistant (PDA).
- To address the concern of multiple “identities” of the same person being created every identity verification goes through a centralized system that monitors for anomalies. Van Oorschot and Stubbebine (2004) refer to this property as “entity uniqueness”
• At transaction time, the location of the transaction is matched to the location of the person (heartbeat locator) to create a real time, online verification system

Conclusion

This report gave a brief introduction to the phishing problem and background information on how an attack is executed. A selection of countermeasures was then discussed: - the EarthLink and eBay browser plugins as viable client side defences; scanning and alerting software was also looked at as an alternative to user intervention and email verification tools to counter spam. The section on website verification solutions looked at the current attack vectors that are being used to create fake websites and their countermeasures. There was also a discussion on how to guarantee the legitimacy of messages and their source.

Two factor authentication was described as well as a commercially available implementation of this idea. An argument against the effectiveness of two factor authentication was also provided. Lastly there was brief discussion of identity theft and some proposals on how to handle this growing problem.

Technology alone will not stop what has become a lucrative activity. Greater awareness is needed and users must be educated about the risks associated with divulging personal information online. Businesses must also protect the user’s information more securely for example by ensuring that information is never transmitted in clear text.

There is also a need for greater co-operation and information sharing between financial institutions, government agencies and technology providers. And if we ever want to provide a deterrent to this activity then we need better trained law enforcement agents to find and prosecute the perpetrators of these crimes.
In conclusion one can see that there are many attempts to propose and implement solutions to the phishing phenomenon. However the problem domain is constantly changing, as new solutions are found, the attack vectors are modified to become more sophisticated.

There is a long road ahead before the tide of phishing attacks is stemmed, if indeed that is possible.
References


Drake, Christine, E., Oliver, J., Koontz, Eugene, J. “Anatomy of a Phishing Email”. 2004


Kerner, S.M. “The Cost of Phishing Hits $1.2 Billion”, 6 May 2004
www.internetnews.com/ec-news/article.php/3350891


Morrison, Scott., Nuttall Chris. “Gone phishing Fraud based on bogus e-mail is causing concern among banks and online retailers” (FT Report – FT-IT Review). 23 June 2004


Rudd, Barney. “An analysis of Phishing and possible mitigation strategies”. 27 June 2004


Tumbleweed Communications. “Using Digital Signatures to Secure Email and Stop Phishing Attacks (A Tumbleweed Whitepaper)”. 2004

Tynan, Dan 2004, “EarthLink Readies Anti-Phishing Tool ScamBlocker is designed to track, block access to known swindle sites”, PC World, viewed 29 October 2005, <http://yahoo.pcworld.com/yahoo/article/0,aid,115652,00.asp>
