Can Identity Theft Defense be Practically Effective? A TAM-Derived Survey of Software-Based Deterrence to Phishing and Pharming

Charles McClain, Capella University

ABSTRACT
Phishing uses technical legerdemain and social engineering to defraud on-line consumers into divulging confidential information. Pharming is a complimentary activity which misdirects clueless internet users to phony destinations, where they are induced to voluntarily give up their valuable data through Pharming schemes. These crimes are on the rise. Filings with the Anti-Phishing Working Group (APWG) indicate that, on average, 456 incidents of Phishing were occurring each day during the month of July, 2005, and these were only those incidents which were reported to APWG. This paper discusses practical challenges to the problem of identity theft prevention through an evaluation of the perceived user acceptance over a variety of software products which share the goal of enabling web-users to avoid becoming ID Theft victims, based on hypothesis of the Davis Technology Assessment Model (TAM).

INTRODUCTION
Fraudulent misappropriation of identity has been a popular method of criminals to steal unsuspecting victims’ persona and resources, in both fiction and reality, since long before the internet or the personal computer. In Frederick Forsythe’s “Day of the Jackal,” we recall that the Jackal - whose real identity we never learned - in an attempt to assassinate General DeGaulle, stole two identities, one to obtain an actual birth certificate and the other to acquire the Danish passport of a teacher, the physical appearance of whom the Jackal duplicated by disguise. Another classic ploy involves the criminal dressing up like a policeman to fool the hapless into willingly following nefarious instructions. The games remain the same, only the toys are new.

In our cyberworld, Phishing and Pharming are terms-of-art describing criminals’ use of a combination of technology and social engineering to fool users into revealing sensitive information enabling the perpetrator to fraudulently access the victim’s financial resources. In terms of social contrivance, advanced phishing examples misappropriate the “look and feel” of a familiar or respected company website, inducing a trusting user to reveal sensitive data, such as social security or credit card number, passwords, account usernames, and the like. Technically, email may be used to surreptitiously install “malware,” such as Trojan keyloggers and other spyware, on an unsuspecting customer’s computer which collects sensitive credentials for later criminal misuse. This commonly occurs through the use of spoofed web pages or on-line forms. Pharming employs such tools, often through Domain Name Service (DNS) hijacking or poisoning, to fraudulently redirect unsuspecting users to fake web sites or proxy servers.

The incidence of such attacks is has increased at an alarming rate over the past year. An international security organization, the Anti-Phishing Working Group (APWG), tracks reports of such activity and has identified a disturbing trend portrayed in Figure 1.

While we see a temporary decline in these figures since June, 2005, the persistent trend is upward, in effect doubling in monthly occurrence over the term represented in Table 1. It is also important that the existence of this organization has not been widely reported and thus we can assume that the number of Phishing occurrences may, actually, be substantially higher than indicated in the APWG data. Interpretive data mining further illuminates the current state of the Phishing problem, as shown in Table 1.

The problem of identity theft through Phishing and Pharming has increased in proximate relation to the expansion of Internet commerce. Kerner reports that in February, 2005, the FTC announced that for the fifth consecutive year, identity theft constituted the most frequent class of complaints of fraud filed with the FTC. With increasing frequency, individual states are taking the initiative and holding government/industry seminars on finding solutions to identity theft, such as that which was hosted by California Governor Arnold Schwarzenegger in March, 2005.
PHISHING AND PHARMING IN DETAIL

As noted in Wikipedia, the odd spelling of these terms relates to their origin amongst hackers early in the history of personal computers and the internet, when hackers began to substitute the letters “ph” for “f” in recognition of the early roots of their work related to manipulation of the tone signals of telephones or “phones.” The acknowledgment is greater than semantic, in that it implies acquiescence to the culture which evolved, at that time, and which, from the beginning, was devoted to fraudulent activities.

Generally speaking, Phishing and Pharming have evolved to focus primarily on financial institutions, such as banks and credit unions, or companies with primarily concerned with facilitating cash transactions, such as credit card companies or on-line auction services, such as eBay. Of 119 reports filed with the APWG between July, 2004 and July, 2005, all but 13 were in these categories, and the remainder of these reports referred to firms which have high cash flows in the regular course of their business.

The most common type of Phishing attack, and the earliest examples of this activity, involves the sending of an email representing a legitimate business activity and requesting sensitive information from the recipient. This has come to be called “spoofing.” Typically, in such an attack, the user is requested to “verify” sensitive information by entering the information in response. Initially, this information was expropriated via a return email to the requestor. With the rise of cheap and easy-to-use web graphics design programs, the technique became more sophisticated, when the perpetrators began to expropriate web pages from legitimate business sites in their entirety, and requesting the entry of the purloined information in official-looking web forms. Everything about these sites, including security certification can be made to appear legitimate. Another form of Phishing is called an “exploit-based” attack, which occurs by exploiting a known defect in browser programs, such as Microsoft’s Internet Explorer or Mozilla Firefox, to install programs, such as keyloggers, via viruses contained in emails or websites. This type of Phishing is collateral to the emphasis of this paper, and is best deterred by making sure that security patches provided by software publishers, e.g. Microsoft and www.mozilla.org. It is noteworthy that fraudster’s targeted information is much more difficult to acquire with this method, as data discrimination is far more complex. Recently, a derivation of the email-based Phishing, called “Spear Phishing,” has evolved which uses surreptitiously obtained customer lists for well-known companies to specifically target those customers with increasingly sophisticated scams. Kerstein (2005) has estimated that by the end of 2005, annual business losses resulting from Phishing fraud have grown to over $2 billion.

Pharming exploits vulnerabilities in DNS server software, which resolves internet site names and IP addresses, to allow the hijacking of the domain name for a specific web site and secretly re-direct traffic to another location where Phishing is used to fool the user into believing the site is legitimate. A common example is the replication of a customer’s banking web site, where it seems innocuous for the customer to enter the confidential information sought by the criminals.

ANTI-PHISHING EFFORTS

A great deal of public focus has been given to the issue of identity theft in the past two years. A cynic might opine that this attention was accelerated by the fact that in February, 2005, the identities of several U.S. Senators were included in a large quantity of Bank of America confidential information data tapes which were stolen by airline workers, as reported by Snyder. Be that as it may, The Federal government has taken positive action during 2005, beginning in January, when the Federal Trade Commission (FTC), filed its first action for Phishing against a California teenager, who was charged with misappropriating an AOL web site for the purpose of stealing credit card numbers, as reported by Legon. In March, 2005, legislation was introduced by Senator Patrick Leahy, which would impose 5-year prison sentences and fines for the falsification of corporate email or web sites. In addition to the Federal Government, American companies are also beginning to get serious about identity theft. Also in March, Microsoft Corporation filed 117 “John Doe” lawsuits in Federal Court against individuals for stealing passwords which were used to attempt access to confidential information, as reported in Wikipedia.

In addition to social and governmental action against Phishing, several software companies have introduced products intended to thwart Phishing attacks. In this paper, ten such products were selected as representative of those available in late 2005. These included:

1. Adorons Security Easy©;
2. AntiPhish©;
3. FraudEliminator©;
4. Green Armor Solutions Identity Cues©;
5. Inspector Brown Fraud Buster©;
6. Netcraft Toolbar©;
7. PayPal Safetybar®;
8. Secure Legal Action Management©;
9. SpoofGuard©; and,
10. SpoofStick®.

In a following section, each of these software packages was evaluated for practical effectiveness in inhibition of identity theft, in terms of a common set of criteria intended to reflect relative user acceptance.

USER ACCEPTANCE

Intuitively, it seems reasonable to assert that the effectiveness of anti-Phishing software is critically dependent on whether and to what degree the software is accepted by the internet users to whom it is distributed; Furthermore, browser software, including Internet Explorer and Firefox, are freely distributed to internet users, either independently or as part of an underlying operating system (OS), such as MS Windows; thus it logically follows that an acceptable cost threshold for anti-Phishing would be low, if the use of the software-based anti-Phishing solution is broad enough to have significant impact on the problem of identity theft.

A survey of available research literature indicates that the effectiveness with which a software solution is implemented is vitally related to its acceptance by users. The basic premise of this view was laid out by Davis (1986), in his doctoral dissertation proposing a technology assessment model (TAM) for evaluating the acceptance of new information systems. In this construct, several hypothesis’ were set forth, including user trust of the software source, user perceptions about the utility of the software, and the concept of ease-of-use – or user-friendliness – of the software. A number of follow-up studies have supported this work. Scholefield and Zedan (1992) replicated the TAM research for user acceptance of real-time and fault-tolerant software. Interest in the TAM concept has increased with the dramatic expansion of internet commerce beginning in the late 1990’s. Gefen (2003) augmented to TAM structure with the strengthening of user acceptance of on-line shopping through force of habit. Hsu and Lu (2004) positively applied the TAM to on-line gaming with the addition of social and belief-related influences. Perhaps the most germane extension of this work for our purposes was performed by Soo and Han (2002). In their research, the validity of the TAM was replicated and validated for the field of internet banking, through a broadly based and statistically validated survey which lead to the conclusion that the most critical factor in customer acceptance of this virtual financial service was the trust which the user held for the banking institution where the internet banking was occurring. This dynamic was closely followed and supported by users’ perceptions of the utility of the banks’ web interfaces and their ease-of-use. The analysis also indicates that perceived and real cost, both in time and money, is integral to a concept of ease-of-use, and affective of a customer’s attitude towards this on-line service, which was also
determined to be a key factor in customer’s ultimate intent to use internet banking.

In this paper, the assumption is made, based on the preponderant targeting of financial institutions in reported Phishing schemes, that the TAM and related research relevant to customer acceptance of internet banking has parallel implications for the anti-Phishing software reviewed herein, although the search of available literature indicates that directly related research in this area is not available for review.

ASSUMPTIONS, METHODOLOGY, AND ANALYSIS

In analyzing the relative effectiveness of the various anti-Phishing software solutions, the APWG statistics (Table 1) note that 46% of the reported Phishing sites contained some naming elements of the misappropriated businesses their sites purported to represent. How extensive these similarities were before detection and quarantine by the internet service providers (ISP) providing the platforms is an open question, but given the increased attention of Federal regulators and Congress, it would be surprising if these service providers were oblivious to the issue. Furthermore, the data indicated that another 41% of the fraudulent sites are simply IP addresses, with no implicit relationship to the target company. These factors lead to the conclusion that the most important information which users of anti-Phishing software require is the name of the website to which they’ve been directed. This, in most cases will alert them to the true identity of the web-page provider. Features provided in the software packages reviewed varied greatly, and, while all packages provided the true identity of websites to which users were connected, many provided far more features, such as password authentication, email-based Phishing message trapping, client-resident databases for accumulation of information about Phishing websites, and internet-based data collection and collaborative Phishing information services. Whereas many of these features are ingenious and technically interesting, their utility to consumers of on-line services is questionable. This issue is of particular concern in terms of the financial cost of some of this software, particularly in light of the fact that the underlying mechanism which enables most of these products is a browser which the user acquired freely or as part of an operating system, where the incremental cost of the browser software is minimal or obscured from the user. Other issues which raised questions regarding user acceptance of anti-Phishing software included the fact that some of the packages do not work on both of the most widely used browsers, Microsoft Internet Explorer and Mozilla Firefox. One of the packages, PayPal Safetybar, works within a specific family of email products, Microsoft Outlook, which, while probably used by a majority of the Windows-based computer users, is not the only email client software available. For example, Mozilla.org provides free client email software in both their Thunderbird and Mozilla suite products.

The methodology used in analyzing these software packages is based on review by three users, using the criteria stipulated below. All are experienced computer professionals, one of whom has over twenty years of experience, having used web-based internet access since its inception with the introduction of Version 2 of the Mosaic browser in late 1993, developed at the National Center for Supercomputing Applications (NCSA). The point is that these users were chosen because they are very familiar with web-based technology. This formed the basis of the evaluation, which focused on the ease of use of the various packages reviewed. It was arbitrarily determined that if the software could downloaded, installed, and configured in less than 10 minutes, the software was rated easy to use; if the software took between 10 minutes and 30 minutes to download, install, and configure, it was rated as moderately easy to use; and, if the software took longer than 30 minutes to download, install, and configure, it was rated it as complex in terms of ease-of-use. This reasoning was based on an assumption that most on-line users would have neither the level of skill, experiential predisposition, nor patience to work through product installation and configuration with greater facility than the testers; thus ease-of-use would be unlikely to be greater in the general universe of potential users of this type of software. These preconceptions may be questionable, but that is beyond the scope of this paper.

Another product aspect which was included in this study concerned feature levels, based on a Likert rating scale of 1 through 5. These levels were based on the following criteria:

1. Actual URL Identification of Phishing Site
2. #1 + User Password Hashing and/or Authentication at Phishing Site
3. Isolation of Phishing-based email messages
4. #2 + Client database of identified and suspect Phishing sites
5. #4 + Subscription to on-line anti-Phishing Service

The results of this investigation are presented in Table 2.

Cost of the software packages was an issue which raised some concern, based on the TAM research which was discussed earlier. In some of the cases, the impression was left that the availability of a “free” anti-Phishing tool was an inducement to gain information about the potential consumer of the software, with entry of substantial personal infor-
tion being required to facilitate download of the software; or, referral to a salesperson, who would call back to provide further information about the anti-Phishing software and services offered. Some of these conversations were thinly veiled sales pitches.

The only packages which were rated easy-to-use were SpoofStick from Corestreet and Netcraft Toolbar. Spoofstick was also the simplest in terms of features, presenting only the actual URL of site to which a user is connected.

This functionality is demonstrated in screenshots taken before installation of SpoofStick (Figure 2 below) and after installation (Figure 3 below). A T1 communications line was used for connectivity, thus downloading of the software was of insignificant time. This was true for many of the products reviewed, but the differences occurred when during configuration of the software for use. Those packages which were rated 2 and 4 for features, required configuration and data entry which took longer than 10 minutes in those cases. The PayPal Safetybar product required installation within Microsoft Outlook or Outlook Express, which were not used by the testers, who were unable to complete testing of its installation. Those products rated 5, required significant time and involvement with software staff and internet registration to gain access to full features.

Several of the packages were relatively feature-rich. In some cases, the identity of a Phishing site was based on the content of a database installed on the client machine or an on-line subscription service. The currency of this information was dependent on ongoing identification and contribution by cooperating users. Given the proliferation of these sites and the fact that revelation of a Phishing site depends on some user initially identifying it as such and reporting it, the author was dubious as to how many users would participate in anti-Phishing programs to the degree necessary for widespread effectiveness.

CONCLUSION

The implications of this study seem significant if one accepts the validity of the research regarding the acceptance of on-line software under the hypotheses of the TAM. Ultimately the success of any software solution is based on its reception and acceptance by users. This appears obvious, prima facie, particularly in the on-line environment of financial services vulnerable to Phishing. The experience in this survey leads to a belief that rigorous surveys and analysis of the reactions of users of on-line financial services when supplied with various anti-Phishing software solutions provide a rich opportunity to discern and facilitate the most effective methods of inhibiting Phishing and Pharming.

REFERENCES


Anti-Phishing Organization Overview, retrieve September 28th, 2005 from Anti-Phishing Organization website, www.antiphishing.org


ENDNOTES

1 A variety of furtively installed software programs which monitor a user’s keystrokes to reveal confidential information.
An unauthorized DNS server modification, which re-directs legitimate URL requests to different Web sites without the users' knowledge. In some cases, such new Web sites' URLs may have single different characters in the names that might go unnoticed. The bogus Web sites may be vehicles to purloin confidential information associated with the intended Web sites.

A category of technical subterfuges which trick a DNS server into perceived receipt of authentic information, when in fact, the information is bogus. Once the DNS server has been so "poisoned," this information is generally cached for some period of time, spreading the effect of the attack, such the illegitimate collection of confidential data, to the server's users.

For an interesting history of “Phreaking,” and the involvement of several people who later became computer industry celebrities, see http://en.wikipedia.org/wiki/Phone_phreaking
Related Content

Survey on Privacy Preserving Association Rule Data Mining
[www.irma-international.org/article/survey-on-privacy-preserving-association-rule-data-mining/178163](www.irma-international.org/article/survey-on-privacy-preserving-association-rule-data-mining/178163)

Modeling Stock Market Industrial Sectors as Dynamic Systems and Forecasting
[www.irma-international.org/chapter/modeling-stock-market-industrial-sectors-as-dynamic-systems-and-forecasting/112821](www.irma-international.org/chapter/modeling-stock-market-industrial-sectors-as-dynamic-systems-and-forecasting/112821)

Covering Based Pessimistic Multigranular Approximate Rough Equalities and Their Properties
[www.irma-international.org/article/covering-based-pessimistic-multigranular-approximate-rough-equalities-and-their-properties/190891](www.irma-international.org/article/covering-based-pessimistic-multigranular-approximate-rough-equalities-and-their-properties/190891)

Hybrid Computational Intelligence
[www.irma-international.org/chapter/hybrid-computational-intelligence/112325](www.irma-international.org/chapter/hybrid-computational-intelligence/112325)

Feature Extraction in Content-Based Image Retrieval
[www.irma-international.org/chapter/feature-extraction-in-content-based-image-retrieval/113047](www.irma-international.org/chapter/feature-extraction-in-content-based-image-retrieval/113047)