CHAPTER 1

INTRODUCTION TO PHISHING

Steven Myers

1.1 WHAT IS PHISHING?

> **Phishing:** A form of social engineering in which an attacker, also known as a phisher, attempts to fraudulently retrieve legitimate users' confidential or sensitive credentials by mimicking electronic communications from a trustworthy or public organization in an automated fashion. Such communications are most frequently done through emails that direct users to fraudulent websites that in turn collect the credentials in question. Examples of credentials frequently of interest to phishers are passwords, credit card numbers, and national identification numbers.
>
> The word phishing is an evolution of the word fishing by hackers who frequently replace the letter 'f' with the letters 'ph' in a typed hacker dialect. The word arises from the fact that users, or phish, are *lured* by the mimicked communication to a trap or *hook* that retrieves their confidential information.

In the last few years there has been an alarming trend of an increase in both the number and the sophistication of phishing attacks. As the definition suggests, phishing is a novel cross-breed of social engineering and technical attacks designed to elicit confidential information from a victim. The collected information is then used for a number of nefarious deeds including fraud, identity theft and corporate espionage. The growing frequency and success of these attacks has led a number of researchers and corporations to take the problem seriously. They have attempted to address it by considering new countermeasures and researching new and novel techniques to prevent phishing. In some cases the researchers have suggested old and proven techniques whose use has fallen out of favor, or were considered outdated. In other cases, new approaches are being developed. In this book an

Phishing and Countermeasures, Edited by Markus Jakobsson and Steven Myers

overview of current and likely future evolutions of phishing attacks will be given. Additionally, an overview is given of the security countermeasures that are being developed to counter them. We have tried to take an all encompassing approach, looking at technologies with diverse backgrounds, from technical solutions aimed at directly halting the attacks; to the likely effects of legislation and social networks which aim to make the attacks more risky or easily identifiable, and therefore less profitable for phishers. Additionally, since phishing attacks are often successful because the average user does not understand, and thus cannot make use of, many of the currently available and deployed security mechanisms, we look at human-centered solutions and approaches, that consider a more holistic view of security that includes the user. This approach accepts the fact that average users do things that security experts would rather they didn't, such as using the same password at multiple sites, or choosing their dogs' names as passwords for their online bank accounts.

In the remainder of this chapter the problem of phishing is introduced for the uninitiated. It begins with a brief history of phishing, and then continues on to motivate the need to study the problem by discussing the different costs associated with the attack. Next, the typical anatomy, tools and techniques of current phishing attacks are covered, including two relatively straightforward examples of actual phishing attacks. Lastly, phishing attacks have constantly evolved, and will continue to do so. Therefore, the natural evolution of phishing and the difficulty of protecting users from these attacks are briefly discussed. Hopefully, by this point the reader will realize the potential problems that phishing may realize in the future, and be well motivated to continue further into the book which discusses both expected evolutions of phishing, and potential countermeasures.

1.2 A BRIEF HISTORY OF PHISHING

Phishing originated in the early 1990's on the America Online (AOL) network systems. At the time many hackers would create false AOL user accounts, by registering with a fake identity and providing an automatically generated, fraudulent credit card number. While these credit card numbers did not correspond to actual credit-cards nor the made up identity, they would pass the simple validity tests on the credit card numbers that were performed by AOL (and other merchants at the time), leaving AOL to believe that they were legitimate. Thus, AOL would activate the accounts. The AOL accounts that resulted from such attacks would then be used to access AOL resources at no cost nor risk to the hacker; the account would remain active until AOL actually tried billing the credit card associated with fraudulent account, determined it was invalid, and deactivated the account. While such attacks should not be considered phishing, AOL's response to these attacks would lead hackers to develop phishing.

By the mid 1990's AOL took proactive measures to prevent the previously mentioned attack from taking place by immediately verifying the legitimacy of credit card numbers and the associated billing identity. This resulted in hackers changing their method of acquiring AOL accounts. Instead of creating new accounts with fraudulent billing information linked to made-up identities, phishers would steal the legitimate accounts of other users. In order to do this, phishers would pose as employees of AOL and contact legitimate AOL users, requesting their password. Of course users are not completely naïve and asking directly for a password was unlikely to result in it being given out, but the phishers would provide the users with a legitimate-sounding story that entice them into providing the information. For instance, a phisher would often contact a user and inform them that for security purposes they needed to verify the user's password and would thus need to provide it. This contact would

generally be initiated through email or through AOL's Instant Messaging service, but the email and instant messages would be "spoofed" to appear to come from an AOL employee. Because of the legitimate sounding reason and the appearance that the request for the password came from an authoritative source, many users willfully gave up their passwords to the phishers. The phishers would then use the accounts for their own purposes, accessing different billed portions of AOL's site, with the charges being billed to the legitimate account holder.

Attacks of the form just described are probably the first example of true phishing attacks. Phishers would use social engineering and identity impersonation through spoofing to steal legitimate users' passwords for fraudulent purposes. Further, by this point, the term phishing had already been coined to describe such attacks. The quote below, from a hacker posting on an early Usenet newsgroup `alt.2600`, is one of the earliest known citations on phishing and references the type of attack just described.

> It used to be that you could make a fake account on AOL so long as you had a credit card generator. However, AOL became smart. Now they verify every card with a bank after it is typed in. Does anyone know of a way to get an account other than phishing?
>
> mk590, "AOL for free?," alt.2600, January 28, 1996

Based on the relative success of these attacks, phishers have slowly been evolving and perfecting their attacks. Phishers no longer limit their victims to AOL's users, but will attack any Internet user. Similarly, phishers no longer restrict themselves to impersonating AOL (or agents thereof), but actively impersonate a large number of online e-commerce and financial institutions. Finally, the goal of phishers tends to be more ambitious. No longer do they satisfy themselves with hijacking a user's online account in order to get free access to online services. Rather, they actively attempt to obtain valid credit-card numbers, bank account details, Social Security and other national identification numbers, for the purposes of theft, fraud and money-laundering; and targeted attacks on employees' usernames and passwords are done for the purposes of corporate espionage and related criminal activities.

Finally, the phisher should not be viewed as the lonely high-school or college student who hacks away from his bedroom in the evening. There are reports that organized crime is actively organizing phishers, in order to profit off of fraud and to engage in money laundering services. There are also fears that terrorists may be obtaining funding through similar means. Further, phishing represents a true free market economy. A study done by Christopher Abad [1], which involved monitoring chat rooms that phishers were inhabiting, shows that phishing is not generally done by *one* individual, but rather there has been a specialization of labor, allowing different hackers, phishers, and spammers to optimize their attacks. Abad found that there were several categories of labor specialization such as mailers, collectors, and cashers, as defined below:

Mailers are spammers or hackers who have the ability to send out a large number of fraudulent emails. This is generally done through bot-nets. A bot-net consist of a large numbers --often in the thousands-- of computers that have been compromised, and which can be controlled by the mailer (bot-nets are also frequently referred to as zombie nets). For a price, a mailer will spam a large number of inboxes with a fraudulent email directing users to a phishing website.

Collectors are hackers who have set up the fraudulent websites to which users are directed by the fraudulent spam, and which actively prompt users to provide confidential information such as their usernames, passwords, and credit-card numbers. These websites are generally hosted on compromised machines on the Internet. Note that

collectors are frequently customers of mailers, in that they will set up a fraudulent site and then pay a mailer to spam users in the hopes of directly a large number of victims to the site.

Cashers take the confidential information collected by the collectors, and use it to achieve a pay-out. This can be done in several manners, from creating fraudulent credit-cards and bank cards which are used to directly withdraw money from automated teller machines to the purchase and sale of goods. Cashers are known to either pay collectors directly for the personal information corresponding to users or charge commission rates, where they receive a certain percentage of any funds that are eventually retrieved from the information. The price paid or the commission rates charged are dependent on the quality and amount of data provided and the ability of casher to attack and defraud institutions and service providers related to the collected user account information.

1.3 THE COSTS TO SOCIETY OF PHISHING

Determining the exact cost of phishing is clearly a non-trivial task, as the phishing economy is clearly a black market that does not advertise its successes. Further, many users, corporations and service providers are unwilling to acknowledge when they have become the victims or the targets of a phishing attack due to fear of humiliation, financial loses, or legal liability. However, there is more to the costs of phishing than just the cost of the fraud. Specifically, there are three types of costs to consider: *direct*, *indirect*, and *opportunity* costs:

Direct costs are those that are incurred directly because of the fraud. In other words it compromises the total value money and goods that are directly stolen through phishing. There have been several groups that have attempted to put a value on the direct costs of phishing, and while the results have been fairly inconsistent they have all shown that the direct costs alone are staggering. The Gartner Group [6] valued the direct costs of phishing fraud to U.S. Banks and credit-card companies alone at $1.2 billion for the year of 2004, whereas TRUSTe, a non-profit privacy group, and the Ponemon Institute [12], a think-tank concerned about information management issues, pegged the value at $500 million for losses in the United States. Meanwhile, the Tower Group [11] pegged the losses at $150 million world wide. Clearly, even the lowest figures represent extraordinary costs that need to be dealt with. While the above numbers are all estimates of actual phishing, in March of 2004 a phisher, Zachary Hill of Houston, pleaded guilty to phishing , and the facts that emerged from this case provided some preliminary concrete numbers. The FTC claims he had defrauded 400 users out of at least $75,000, for an averages of $187.50 per victim.

Indirect costs are those costs that are incurred by people because they have to deal with and respond to successful phishing attacks, but these costs do not represent money or goods that have actually been stolen. Examples of indirect costs for service providers include the costs associated with customer service and call centers that service providers must deal with, when they are targeted by attacks, the costs of resetting peoples passwords, temporarily freezing their accounts, etc. Examples of indirect costs for victims include the time and money expended in reclaiming one's identity after identity theft, tracking down fraudulent charges and having credit-card

companies excluding them from the bill, and dealing with credit-rating agencies to ensure that their credit rating is not devastated by an attack.

The website Bank Technology News reported [5] in April of 2004 that one of the top 20 U.S. banks had to field 90,000 phone calls per hour for five hours after a phishing attack in February. In the same article David Jevans, chairman of the Anti-Phishing Working Group, estimates that phishing costs financial institutions about $100,000 to $150,000 in brand devaluation per phishing attack. An article in CSO Online [2] states that the costs to Earthlink, a large U.S. Internet Service Provider (ISP), for helping phishing victims deal with the attacks repercussions, such as reseting passwords, are approximately $40,000 per attack. Additionally, at that time Earthlink was dealing with approximately eight unique attacks per month.

Opportunity costs are those costs that are associated with forgone opportunity because people refuse to use online services because of the fear of phishing, or are otherwise suspicious of them. For instance, if a user is too afraid to use online banking because of the fear of phishing, then the opportunity costs to the bank is the difference in the many different costs associated with dealing with that customer online versus offline. Similarly, the same customer has the lost the ability to perform many banking options from the comfort of their home, such as paying bills, get account updates, etc..., and so the customer must now travel to a bank or ATM machine to perform the same service. The opportunity costs associated with users losing confidence in e-commerce is especially high for those merchants that only sell online, such as Amazon.com. Such merchants lose potential customers, as they have no corresponding brick and mortar presence. In a 2005 survey by the Gartner Group [7], it was found that 28% of online banking customers said their online banking activity was influenced by online attacks such as phishing. In particular, 4% of customers had stopped paying bills online and 1% of customers had stopped banking online altogether because of online attacks. Similarly, the Gartner Group estimates that a financial institution saves 45 cents every time that a statement is emailed as opposed to it being sent physically in the mail. Therefore, if users stop trusting email from their financial institutions, a large bank could easily have an opportunity to save millions of dollars per year just in potential postal expenses!

1.4 A TYPICAL PHISHING ATTACK

Currently, the most common form of phishing attacks include three key components: the lure, the hook, and the catch. They are as described below.

The Lure consists of a phisher spamming a large number of users with an email message that typically, in a convincing way, appears to be from some legitimate institution that has a presence on the Internet. The message often uses a convincing story to encourage the user to follow a URL hyperlink encoded in the email to a website controlled by the phisher and to provide it with certain requested information. The social engineering aspect of the attack normally makes itself known in the lure, as the spam gives some legitimate sounding reason for the user to supply confidential information to the website that is hyperlinked by the spam.

The Hook typically consists of a website that mimics the appearance and feel of that of a legitimate target institution. In particular, the site is designed to be as indistinguishable from the target's as possible. The purpose of the hook is for victims to be

directed to it via the lure portion of the attack and for the victims to disclose confidential information to the site. Examples of the type of confidential information that is often harvested include: usernames, passwords, social-security numbers in the U.S. (or other national ID numbers in other parts of the world), billing addresses, checking account numbers, and credit-card numbers. The hook website is generally designed both to convince the victim of its legitimacy and to encourage the victim to provide confidential information to it with as little suspicion on the victim's part as possible.

The Catch is the third portion of the phishing attack, which some alternatively call the kill. It involves the phisher or a cashier making use of the collected information for some nefarious purpose such as fraud or identity theft.

A more precise and detailed view of the different components of a generalized phishing attack will be considered in Chapter 2, but the three components just listed will suffice to describe the most basic and common phishing attacks. In the next subsections, several real-world examples of recent phishing attacks will be given, so that readers who are unfamiliar with the concept are exposed to some concrete examples. Readers who are already familiar with what a typical phishing attack consists of may wish to skip ahead of these subsections.

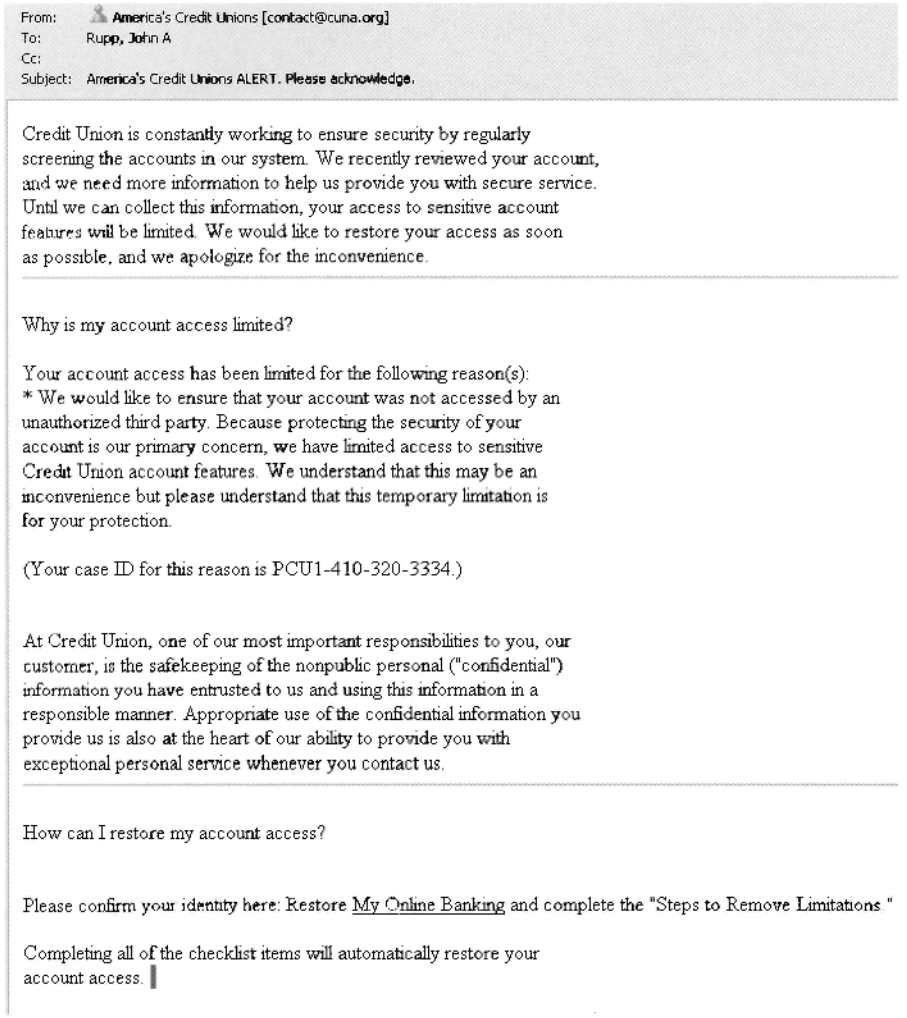
1.4.1 Phishing Example: America's Credit Unions

This example presents a step-by-step walk through of a recent phishing attack that was directed at a user known to the chapter's author. To begin with, the user received the email message depicted in Figure 1.1, on page 7, in his email inbox one morning.

It claims to be a message from America's Credit Unions and informs the user that there is a possibility that his account has been used by a third party and therefore the union has temporarily restricted the ability of his account to perform sensitive operations until such time as authenticating information can be provided. Fortunately, a handy link is provided that will direct the user to the union's web page so he may restore his account.

The email represents the *lure* portion of a phishing attack. Note that the user who received this message does not have an account with America's Credit Unions, but assuming otherwise, then a quick reading of the message would suggest several things: first, that the Union has acted in the user's best interest by temporarily reducing the functionality of the account because they fear it has been attacked; and second, that they have provided the user with a quick and easy solution for restoring his account's functionality. An additional property of the email that make the lure convincing are the fact that the email message appears to be sent from the address `contact@cuna.org`. The email has not actually been sent from this address, but rather through the technique known as *email spoofing* the email message has been made to appear that it was sent by this address. The topic of email spoofing is discussed in more detail in Chapter 3.

By clicking on the link in the lure email, the user is brought to the web page depicted in Figure 1.2, on page 8. This web page represents the *hook* portion of the phishing attack and is designed to acquire confidential information from the user. In order to convince the user to supply the web page with confidential information, it is designed to look very authentic, and clearly has the same look-and-feel as the authentic one it is trying to mimic (the authentic web page of America's Credit Unions is depicted in Figure 1.3, page 9, for comparison purposes). The notion of making a web page that imitates another legitimate page is known as web spoofing, and techniques for doing this will also be discussed in Chapter 3.



From: America's Credit Unions [contact@cuna.org]
To: Rupp, John A
Cc:
Subject: America's Credit Unions ALERT. Please acknowledge.

Credit Union is constantly working to ensure security by regularly screening the accounts in our system. We recently reviewed your account, and we need more information to help us provide you with secure service. Until we can collect this information, your access to sensitive account features will be limited. We would like to restore your access as soon as possible, and we apologize for the inconvenience.

Why is my account access limited?

Your account access has been limited for the following reason(s):
* We would like to ensure that your account was not accessed by an unauthorized third party. Because protecting the security of your account is our primary concern, we have limited access to sensitive Credit Union account features. We understand that this may be an inconvenience but please understand that this temporary limitation is for your protection.

(Your case ID for this reason is PCU1-410-320-3334.)

At Credit Union, one of our most important responsibilities to you, our customer, is the safekeeping of the nonpublic personal ("confidential") information you have entrusted to us and using this information in a responsible manner. Appropriate use of the confidential information you provide us is also at the heart of our ability to provide you with exceptional personal service whenever you contact us.

How can I restore my account access?

Please confirm your identity here: Restore My Online Banking and complete the "Steps to Remove Limitations."

Completing all of the checklist items will automatically restore your account access.

Figure 1.1 A screen shot of a phishing email designed to entice users to visit a phishing website.

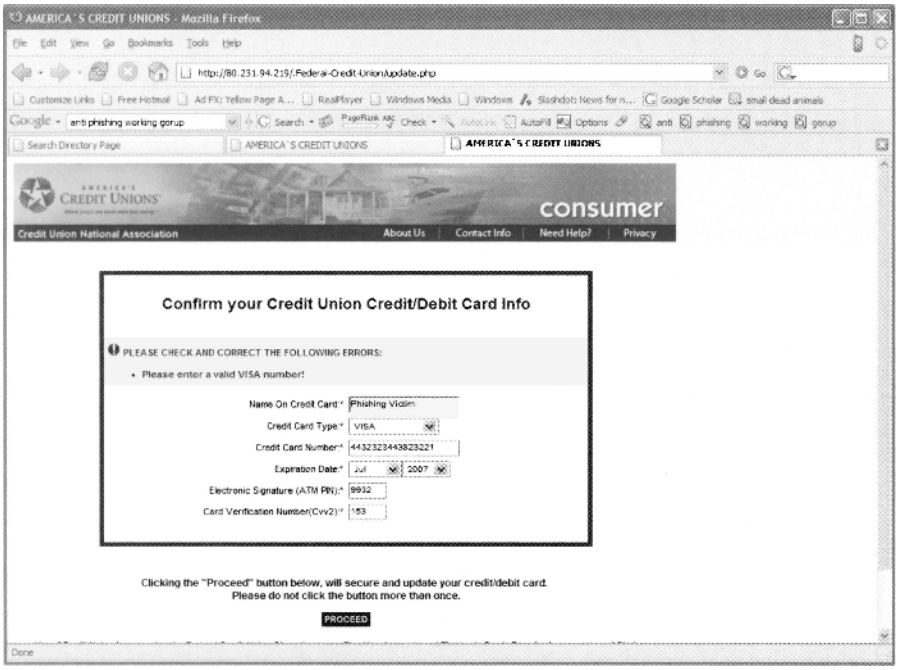


Figure 1.2 A screen shot of the web page that a user is brought to by the phishing email depicted in Figure 1.1. Observe that the web page is designed to actually perform some input validation: a reasonable looking but made up 16-digit credit card number has been entered into the system, but rejected by the web page. There are a number of algorithms available that can be used to check the plausible validity of credit-card numbers, and the author suspects such an algorithm is being used here. The phisher's goal in using such an algorithm is to ensure that people do not inadvertently enter the wrong number.

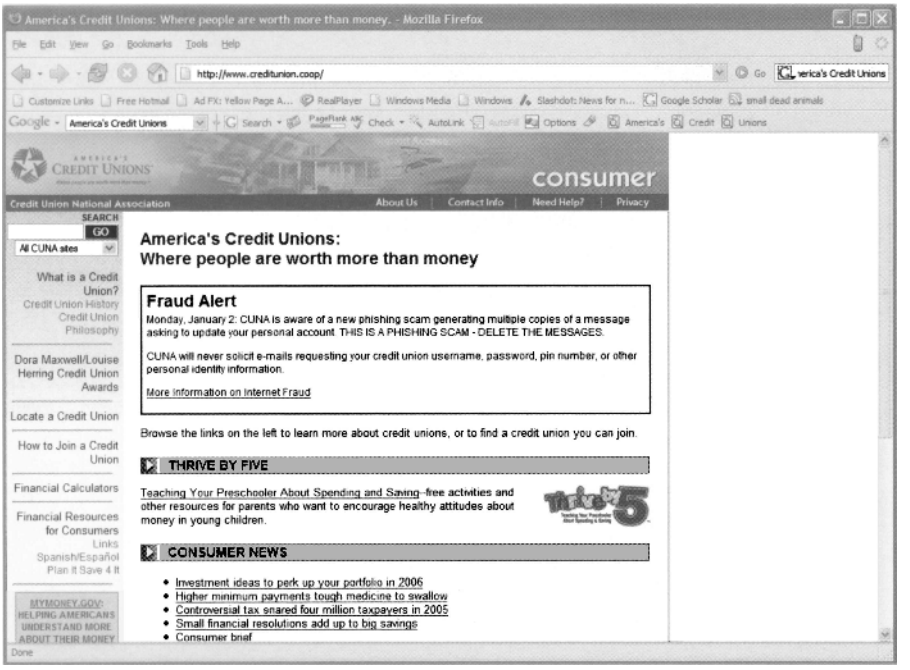


Figure 1.3 A screen shot of America's Credit Unions' actual web page. Notice that the site's look and feel has been effectively duplicated by the phishing site in Figure 1.2. Also note that the web page clearly states that they are aware of a current phishing attack and are asking customers to delete all messages that purport to be from the Union and that ask for personal information to be updated.

The fraudulent *hook* web page that users are directed to by the phishing lure requests certain pieces of "authenticating" information. The requested information includes everything that is necessary to fraudulently purchase items online and to possibly produce a fake credit card to perform purchases in the bricks and mortar world (phishers are able to produce duplicate real-world credit cards for some financial institutions based on the information requested from this site. These cards can then be used to purchase goods, or to buy wire transfers of funds to foreign countries).

Once a user provides the requested documents, the user is given another web page asking the user to confirm the personal identification number (PIN) associated with the credit card that was just entered. This web page does two things: (1) It increases a naïve user's confidence in the page, as users are accustomed to being asked to provide duplicate entries of passwords to ensure they were entered correctly; (2) it reduces the probability that a phisher retrieves an incorrect PIN due to a typo.

After reentering his PIN, the user is brought to the final web page in the attack. It thanks him for entering his authenticating information and informs him that he will be contacted once the information has been confirmed. A victim of this scam, upon reading this web page, is likely to think that there are no problems and go about his business. Note that if at a later point he logs on to his account, then he will have full access to his account (as there

was never any true suspension of privileges), and therefore he is likely to believe the Union has simply re-enabled the account.

The final phase of the phishing attack, the *catch*, does not occur until the credentials that are phished during the attack are fraudulently used. Typically, the phished information will be used to purchase online goods, perform cash advances, or permit wire transfers. As previously mentioned, the *casher*, the criminal who performs the fraud with the gathered credentials, is generally not the same individual as the phisher who acquired the credentials. Further, cashers are generally well-versed on the different techniques for cashing out credentials in manners that ensure the probability that they will be caught by authorities is very small.

1.4.2 Phishing Example: PayPal

In the second example of a phishing attack, an attack on PayPal is considered. PayPal is an escrow payment company most frequently used to pay for goods on the Internet, especially on the eBay auction site (PayPal was in fact used by so many eBay customers that eBay purchased PayPal). In Figure 1.4 we see the *lure* portion of a phishing attempt on PayPal account holders. It suggests that there have been several attempts to access the user's PayPal account from foreign countries, and if the user has not been using the site while traveling, then he or she should visit the PayPal website and verify his or her identity. This verification, it is claimed, will ensure that others cannot access the account.

Further inspection of the lure shows that the phisher goes to the trouble of warning the user about not sharing his or her password with anyone and properly protecting it. Further, the statement gives what is reasonably good advice: to protect themselves from fraudulent websites the user should always open a new browser window, type in the URL `www.paypal.com`, and only then enter his or her credentials. Ironically, if a user were to follow this advice, then the current phishing attack would not effect her. The reason the phisher includes this message in the lure is to make the message seem all the more authentic: Surely an attacker is not going to give security advice to the users being attacked! Of course, the effect works to the phishers advantage, as many users will take the advice to heart, but still click on the link provided in the email claiming to be from PayPal, as clearly PayPal is thought to be trusted.

Other features of this message that make it seem official are that the sender of the email is listed as coming from PayPal (the sender is shown to be `service@paypal.com`) and that the URLs that the user is directed to by the links in the email appear to be legitimate. In particular, we note that the user appears to be directed by the link to the legitimate URL `https://www.paypal.com/us/cgi-bin/webscr?cmd=login-run`, but in reality the link is to the illegitimate URL `http://210.54.89.184/.PayPal/cgi-bin/webscr.cmd_login.php`, which is a web page that corresponds to the *hook* portion of the phishing attack. Observe that to experienced or technical users, the differences between the two URLs are clear, but the average lay-user will either find the two indistinguishable or have trouble distinguishing between them. In the latter case, even if the user can distinguish between the two, she may not be able to identify the legitimate URL.

If a user follows the link provided in the lure, then the browser is launched and the web page depicted in Figure 1.5 is retrieved for the user. This page represents the *hook* of the attack. This first page in the hook directs the user to provide her PayPal username and password to authenticate to the system. Again, the look-and-feel of the actual PayPal website is appropriately spoofed. This can clearly be seen by comparing Figure 1.5 with Figure 1.8, where the latter figure depicts the actual PayPal website. Once the user provides

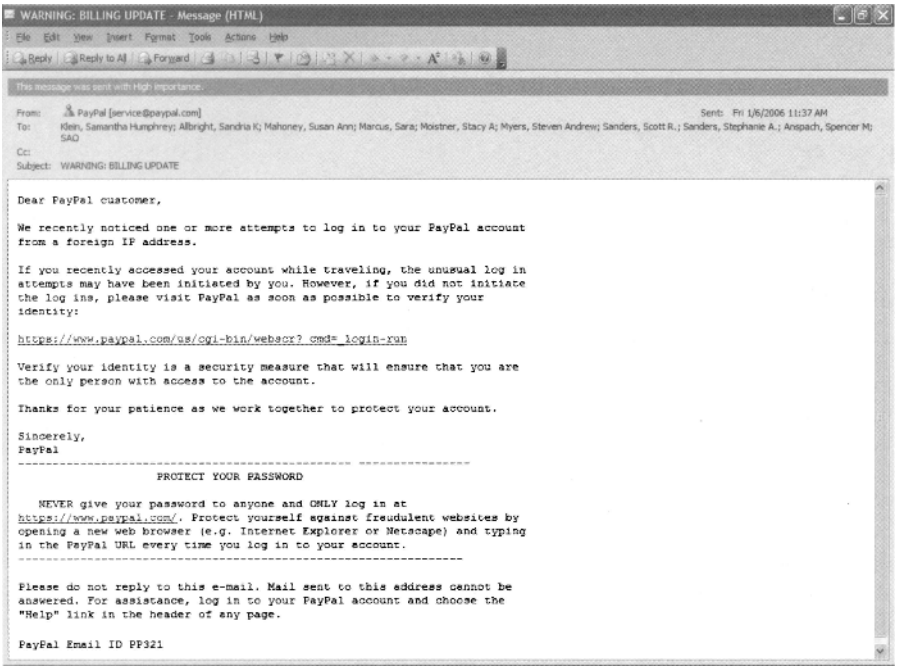


Figure 1.4 A screen shot of the email lure sent in a phishing attack on PayPal. Note that the link to `https://www.paypal.com/us/cgi-bin/webscr?cmd=login-run`, while looking like a link to a legitimate address, actually links to a phishing site.

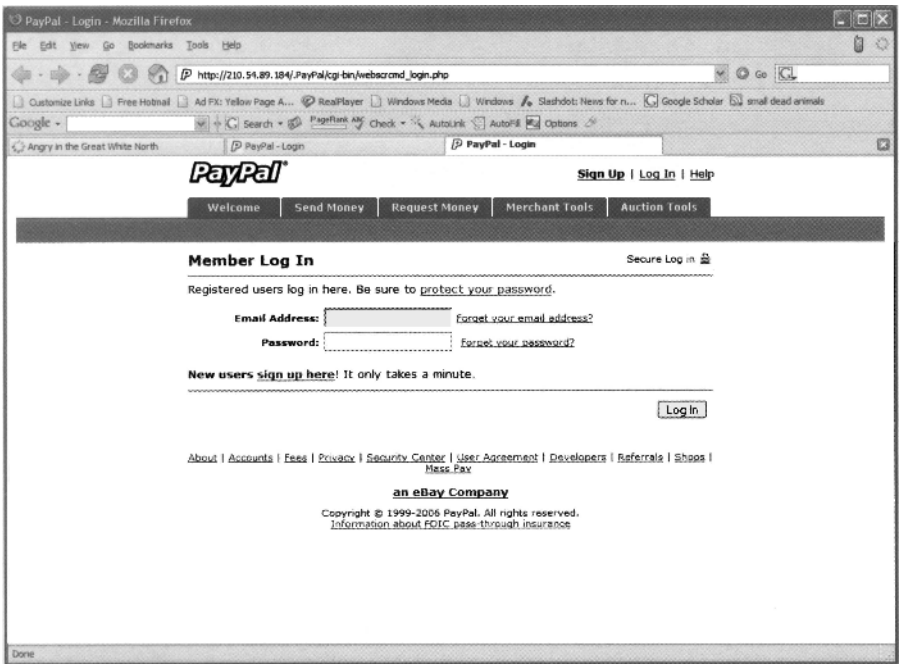


Figure 1.5 A screen shot of the first web page involved in the PayPal Phishing Attack. This screen is meant to retrive PayPal usernames and passwords. Note the similarity between the this website, and the legitimate PayPal website depicted later in Figure 1.8

her username and password, she is brought to a second web page that requests her credit-card and billing address information, as depicted in Figures 1.6 and 1.7. Once the phishing victim has provided this information, they are "logged out" of PayPal and brought back to the *actual* log-in web page for PayPal, as depicted in 1.4.2.

1.4.3 Making the Lure Convincing

In order to convince a potential victim of the legitimacy of their forged email, phishers use several common tricks. We divide these into two categories: social and technical. The social category includes stories, scenarios, and methodologies that can be used to produce a convincing social engineering context that is likely to make a user willing, if not enthusiastic, to provide her confidential information to the phisher. The technical category includes technical tricks that effect the email reader or browser to help support the artificial social context being presented to the user.

1.4.3.1 Social Engineering Methodologies In order to have victims follow the lure of the phishing attack, the phisher must provide a plausible reason or incentive for the victim to click on the hyperlink provided in the spam and must also provide the information requested at the hook. The number of possible scenarios that a phisher might think-up are

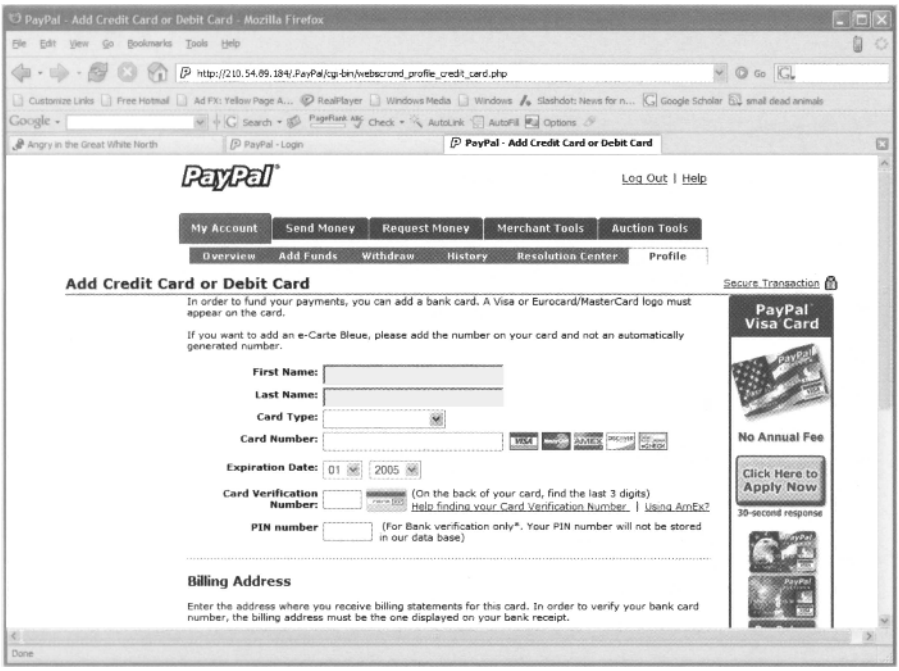


Figure 1.6 A web page in a PayPal phishing attack that requests user's name and credit-card information.

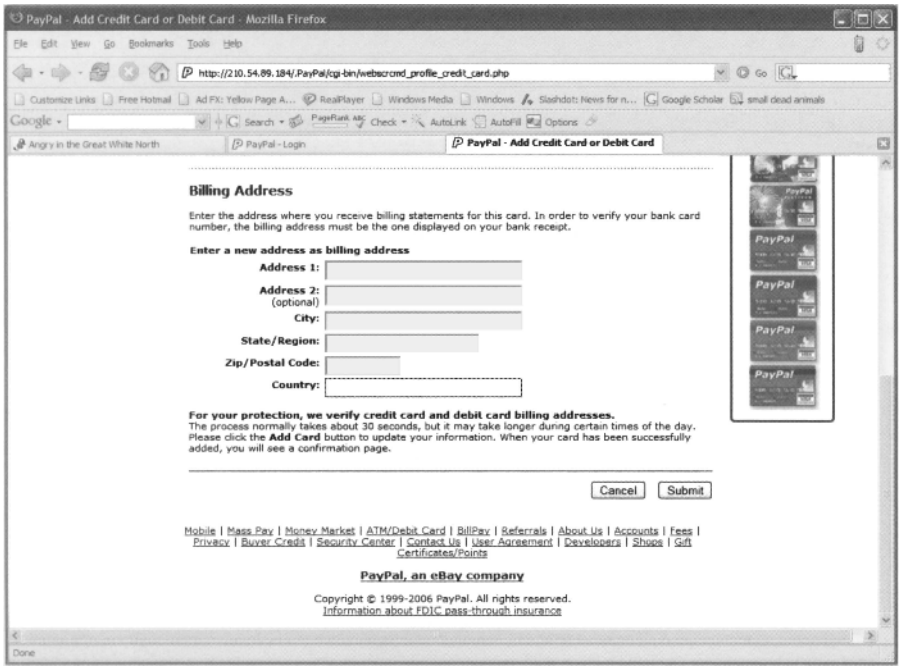


Figure 1.7 A web page in a PayPal phishing attack that requests user's credit-card billing information.

Figure 1.8 PayPal's actual web page. This is provided to compare with the web pages depicted in Figures 1.5, 1.6, and 1.7. Also note that after the phishing attack is completed, the user's web browser brings up this legitimate PayPal web page

almost limitless. Below we enumerate several of the more common scenarios that have previously been used.

Security Upgrade: The users are told that the online service provider is introducing a new service to increase consumer security and protect them from fraud. However, in order for the users to activate or enroll in this new and improved security scheme, they must login to the service provider's site and provide certain types of authenticating information.

Incomplete Account Information: The users are told that their online service provider's account information is out-of-date or missing and that in order to maintain the service the users must log in to the site and update their information. A hyperlink is generally provided in the email that links to the phisher's fraudulent website, and many users follow it in order to ensure that their accounts are not canceled or suspended.

Financial Incentive: The users are told that their online service provider will provide some financial incentive, perhaps in the form of a coupon, discount, a chance to win a prize, or simply a direct cash transfer. The users are told that in order to receive or be eligible for these incentives, they must follow a hyperlink, provide appropriate information, and possibly authenticate to the online service in questions.

False Account Updates: The user is sent a message thanking them for updating their account information at their online service providers website. There is also a warning that if they have not initiated an account update, they should follow a link to the website, log in, and report the fraudulent website. Since the user had not previously updated the information at their service provider (the email is after all sent by a phisher and not the service provider), they are enticed to follow the hyperlink provided by the email, and then authenticate to the phisher's fraudulent website.

While the above list represents many of the common scenarios that are commonly used, it is expected that the social engineering use in phishing attacks will quickly become much more sophisticated. In particular, the scenarios will be customized to match context that a victim is expecting. Such attacks are called Contextually Aware Phishing or Spear Phishing and are discussed in detail in Chapter 6.

1.4.3.2 Technical Tricks A phishing lure also normally uses certain technical tricks that can be used to help reinforce the legitimacy of the email lure and socially engineered story provided within it. Additionally, technical tricks are needed to deliver the lures, so that the fraudulent emails are not traced back to the phisher. Below some of the more commonly used technical tricks are discussed.

Using Legitimate Trademarks, Logos and Images: One of the most obvious but essential technical tricks used by phishers to make a lure convincing is the inclusion of of logos, images, and names that are trademarked or copyrighted by the institution being mimicked. Such logos are regularly used in correspondence and marketing with customers by the institutions being mimicked. The inclusion of such logos gives many users a false sense of security that they're dealing with a legitimate party. Further, many users do not realize the ease with which such images can easily be duplicated in the digital environment, and therefore the mere appearance of legitimacy is sufficient to trick many users into falling for the lure, clicking on the hyperlink, and providing their confidential information to the hook. Phishers have no difficulty in obtaining a

legitimate institution's trademarks and logos, as they can normally be found on the institution's legitimate web site. To add insult to injury, often rather than directly including the image in the lure, a phisher will provide URL links to the legitimate site in the lure email. The result is that sites' legitimate graphics are used in the lure and thus are authentic. In Chapter 3, Section 3.5 a case study is presented that provides legitimate sites a small amount of protection from such attacks.

Email Spoofing: A phishing lure that claims to come from a financial institution appears a lot less authentic if the email address associated with the sender is not associated with the financial institution, but instead some strange email address, such as those addresses associated with bulk free email providers like Hotmail and Yahoo. Therefore, in order to prevent making potential victims skeptical, and potentially spooking them, phishers spoof the email address from which the spammed email appears to come from. That is, they make the apparent sender of the email appear to be different from the actual sender's identity. Both of the lures that were seen in the two phishing examples in Sections 1.4.1 and 1.4.2 used email spoofing to change the apparent identity of the email lure's sender. Email spoofing and countermeasures to prevent it will be discussed in Chapter 3.

URL Hiding, Encoding, and Matching: Some users might become suspicious of a phishing lure if one of its hyperlinks links to a domain name that is clearly incorrect, or not obviously related to the targeted institution. In order to prevent such suspicion on the part of victims, phishers attempt to make the URL that is presented in both email lure and the fraudulent website hook appear official and legitimate.

In the PayPal phishing example, the phisher did both: The link that appears in the email appears to encode for a legitimate PayPal web page, while in actuality it linked to a server controlled by the phisher. But even given that the link went to an incorrect URL, the phisher did his best to make the illegitimate URL appear as though it were a legitimate one. The latter case can be considered a case of URL spoofing. This was done by not using a domain name for the phisher's server, but instead referring directly to the server's IP address. Next, the phisher used a subdirectory entitled `.PayPal`, so that the words PayPal would appear in the URL. The end result is that many lay users do not understand the distinction between a legitimate URL that contains the top-level domain `www.PayPal.com` and another URL that contains an IP address instead of a top-level domain, but contains the words PayPal in the URL. More complicated examples of URL spoofing are discussed in Chapter 3, Section 3.3.

Bot-nets or Zombie-nets: In order to send out a large amount of spammed email that cannot be traced back to the phisher, bot-nets are used. The term bot-nets is short for robot networks. A bot-net consists of a large number of PCs that have either previously been hacked or that have some form of trojan-horse software installed on them, thereby permiting a phisher to control the machines to send a large amount of spam: the machines act as if they were robots under the control of the phisher. By using simple control software, phishers are able to efficiently provide (a) machines in the bot-nets with phishing lures to distribute and (b) lists of email addresses corresponding to potential victims that should receive the spammed email lures. Because the PCs in the bot net are owned by unsuspecting but law-abiding users, tracing the spam back to a sender in a bot-net provides law-enforcement with little to no indication of the phisher's identity, and often no other clues to follow.

Another interchangeable term for bot-nets is zombie-nets. This metaphor comes from the fact that computers in such a network are seen as zombies under control of the phisher, as opposed to robots.

By combining a good socially engineered story, with the technical tricks just mentioned, a phisher is able to construct a luring piece of email that looks very legitimate and authentic to most lay users, and even many experienced users. In fact, MailFrontier, a secure email product provider, has developed a Phishing IQ test[1] that lets people test their ability to distinguish between legitimate communications from major online service providers and phishing lures that have actually been seen in real world attacks; the author has seen many computer security experts fair quite poorly on this quiz.

One key factor that makes it unlikely for potential victims to fall for a phishing lure is if the lure they receive is not directly applicable to them. For instance, the user known to the author who received the lure shown in the America's Credit Union example, from Section 1.4.1, was not fooled by the lure. This was because that particular user did not have an account with a credit union. In other words, the context of the attack was wrong: The lure was written assuming the context that the user had an account with the Union, but he did not. The fact that the contextual information is wrong for most users who receive a given lure is actually one of the best defenses users currently have. Unfortunately, this defense is expected to quickly disappear with the advent of contextually aware or spear phishing attacks. Such attacks are discussed in Chapter 6.

1.4.4 Setting The Hook

The goal of a phisher is to a lure a large number of users to a web page that can be used to collect different types of personal credentials that the phisher will later use for fraud or other nefarious ends. Clearly, a phisher does not want to use his own computer in order to host such a site, as the computer, and thus the phisher, could easily be traced and then possibly convicted. Therefore, in order to set up a web site to act as the hook in a phishing attack, the phisher hacks into another computer or otherwise gets access to a hacked computer and installs the software necessary to run a web site that collects users' information. This generally involves setting up or modifying the setup of web server software such as Apache or Microsoft's Internet Information Server (IIS). Once the web-serving software is installed, the web pages that spoof the appropriate online web service must be created, and scripts must be constructed for appropriately retrieving and processing the data to be retrieved from the forms on these web pages. Once this is done, the *hook* is said to be in place.

Once the *hook* is in place and a significant number of users have fallen victim to the attack by releasing their credentials to it, the phisher must decide what to do with them. Note that if the phisher simply stores them on the hacked computer hosting the website, then the information must be retrieved at a later point, and this presents a problem: If the authorities have traced the computer involved, they may be waiting to trace anyone who logs on to such a machine putting the phisher at risk of capture. Alternatively, if the owner of the hacked machine realizes that his or her computer has been compromised, then she may fix the security vulnerabilities, making the collected information irretrievable. In order to combat these problems, phishers often take the information collected by their web page and broadcasting it directly to different public bulletin boards or Usenet newsgroups, but in order to ensure that no one else can read the collected information, it is encrypted or hidden

[1] At the time of publication, this IQ test could be found at `http://survey.mailfrontier.com/survey-quiztest.html`

with steganographic techniques that make it all but impossible for anyone but the phisher to retrieve the information.

It may sound like it requires a substantial amount of skill to perform everything necessary to develop a web page to act as a *hook* for phishing attacks. Unfortunately, the job is significantly simplified for phishers by the availability of the following technical tools.

Rootkits: These tools are aimed at automating the process of both hacking into and maintaining administrative privileges on an unsuspecting user's computer. The tools take advantage of known security weaknesses or holes in different operating systems and applications that allow them to gain complete administrative control of the user's system. Once such control is established, a rootkit will insert different applications on the hacked machine to make it both easier to control it and prevent the detection of the hack or any of the hacker's movements. This includes hiding the fact that certain processes are running and that certain directories exist, along with cleaning or hiding auditing trails that might alert a system administrator to suspicious activities or resource usages. Rootkits are discussed in more detail in Chapter 4.

While the technical knowledge needed to construct a rootkit is generally quite involved, this does not mean they can only be employed by such knowledgeable individuals. In practice, the kits are constructed by very knowledgeable hackers and then distributed through the Internet's underground. The result is that only a modicum of the knowledge necessary to construct a rootkit that takes advantage of a known security hole is necessary for a phisher to take advantage of it. In fact, phishers are willing to purchase or barter for rootkits that take advantage of the latest security holes, which are less likely to be patched on many systems.

Phishing Kits: Once a phisher has gained administrative privileges on a computer capable of hosting a phishing web site, then the phisher must set up the phishing web server and web site. Note that in order to be useful for hosting a phishing attack, the host computer must have sufficient resources and bandwidth to not only host the phishing site, but to ensure the temporary redeployment of these resources for the phishing attack will go unnoticed. In order to set up the phishing web site, the phisher must install and/or alter the configuration of web server software, develop web pages that look—as much as is possible—like the legitimate web pages they are trying to imitate and spoof, and develop scripts to take information collected from these sites and make it accessible to the phisher.

A phishing kit automates the process of setting up and configuring a web server to host a web site for phishing purposes. Since there are a large number of web sites that a phisher will want to spoof, the web pages that spoof particular companies are generally not included in phishing kits (although some sample web pages might be included for pedagogical purposes, to help the phisher understand how to set up the web pages he will be interested in deploying). The web pages that are needed to imitate a specific online service must be constructed by hand, or for popular online services the phisher may be able to find a corporate schema (these are described next).

Corporate Schemas: Corporate Schemas are sets of web pages that are pre-built to have the look and feel of some legitimate service provider. This allows a phisher to quickly set up web pages on a server that have a look and feel identical to that of some legitimate service provider. Since phishers are interested in attacking many different types of online service providers, it could be time-consuming to be constantly designing web

pages that spoof legitimate ones. Instead of pre-packaging the basic web pages that are necessary to perform the attack, phishers can quickly and easily set up web pages for a number of different online service providers. Further, because these schemas are bought, sold, and traded between phishers, it allows for the specialization of labor, so that those who are best at creating spoofed web pages are encouraged to do so for a number of providers, and the fruits of that labor can be shared by many.

These tools greatly simplify the task of setting up and executing a phishing attack, and therefore, as previously suggested, phishers need not be as technically sophisticated as it might initially appear. Additionally, the amount of time needed to launch an attack is substantially lower than it might be due to the high amount of automation. The end result is that phishing is a crime that offers potentially high rewards, with low associated risks, and increasingly smaller degrees of technical skills are necessary to launch the attacks. This is not a promising outlook!

1.4.5 Making the Hook Convincing

In addition to making the email lure convincing, the phisher needs to make sure that the websites that make up the hook portion of the attack are convincing. Clearly, such a site needs to have the same look-and-feel as the one it is imitating, and if the phisher is able to use a corporate schema, then much of this work is done for him. However, given a good set of imitation web pages, there are still a number of inconsistencies that might lead a user to recognize that they are being fooled, and phishers therefore try to minimize these inconsistencies. Two of the common artifacts on phishing web pages that users might recognize as being odd or out-of-place are discussed next.

Improper URLs: When visiting an online service provider, the URL for the service provider's web page is clearly stated at the top of the web browser. Generally, this URL includes a domain name that corresponds to the online service, and this domain name occurs in the top level of the URL. For instance, in the PayPal phishing example presented in Section 1.4.2, the URL for the legitimate site is `www.Paypal.com`. A phisher needs to use a URL that corresponds to the machine on which the phishing website is hosted. Since this website is, presumably, hosted on a nonlegitimate server, it does not have the same IP address nor the same domain name as the the legitimate server.[2] Therefore, phishers often register domain names that have an appearance similar to that of the legitimate name, and use these URLs for their phishing sites. For example, a phisher might register the domain name `www.paypa1.com`, replacing the letter 'l' with the number '1'. These can be thought of as URL spoofing, or URL homograph attacks, and they discussed in depth in Chapter 3. Because one generally needs to provide contact information to register a domain name, fake or stolen identities are generally used register domain names used in phishing attacks.

Lack of a Secure HTTP Connection: Just about every site that a phisher will mimic employs secure HTTP connections to encrypt and authenticate all information that travels between the users' client computers and the service providers' server computers. When such connections are established, there are certain visual cues that are displayed

[2] Note that it is actually possible for the phisher to make it appear as if the phishing site has the same domain name as the legitimate site, by means of an attack on the DNS infrastructure. Such an attack is called a *pharming* attack and will be addressed further in Chapter 4, but this attack is more advanced and technical than the simple and more technical attacks described in the this introductory chapter.

by web browsers to inform and ensure the user that such secure connections are established. Such security measures were originally implemented to prevent network eavesdroppers from learning confidential information that traversed the Internet, but because the lack of expected visual cues in the browser might alert users to the existence of a phishing attack, phishers would like to either also provide these secure connections or at least have the ability to mimic the visual cues displayed by the browser when such connections are supposed to be established.

In order to establish legitimate secure HTTP connections, a legitimate web server needs to have a cryptographic certificate that allows it to make this secure connection. Further, in order to prevent a user-interface warning from being brought to the users' attention before such secure connections are established, the cryptographic certificates must be acquired from one of only a few numbers of trusted certificate authorities who are implicitly trusted by browser manufacturers (such as Verisign and Thawte). Finally, these certificate authorities are supposed to (and generally do) perform some substantial authentication of the parties to whom they give such commercial certificates, and therefore it is generally considered quite difficult for phishers to get such certificates.

Because phishers have a fair amount of difficulty getting such certificates, they must find ways to achieve the same or similar visual cues and user experiences on their phishing sites in order to deceive users. There have been three common deception methods for mimicking these cues (although the third may not really qualify as a method of deception).

The first method relies on design and security flaws in the browser that let phishers use programming tools such as Javascript to modify the appearance of the browser in order to simulate the visual cues that a secure HTTP connection has been achieved.

The second method of deception takes advantage of users' poor understanding of cryptographic certificates: Phishers construct their own cryptographic certificates, which *have not* been issued by a certificate authority. The phishers use these certificates and attempt to establish legitimate secure HTTP connections between the phishing site and the victim. Because the certificates that phishers use have not been issued by a certificate authority, web browsers warn the user of the potential problems of using such certificates and even go so far as to suggest to the user that they not proceed with this course of action. But, in the end, browsers give users the choice as to whether or not such connections should be established. Given the warning that users receive, one might be tempted to believe they would be hesitant to establish such a connection. Unfortunately this is not the case, as most users have been effectively trained to always accept such certificates. The reasons for this are many, but basically revolve around the fact that many legitimate individuals, institutions, and organizations have made their own certificates in a manner similar to phishers and have required the user to force the browser to accept such certificates for a number of legitimate purposes, and the result is that users rarely read the warning message, nor understand the true potential for harm that comes from accepting such certificates, as they have effectively been trained to ignore the warning. Once a phisher's certificate has been accepted, the browser does not provide any visual cues that distinguish the fact that the certificate in use is not issued by a legitimate certificate authorities. Therefore, once a phisher's certificate is accepted, all visual security cues of a secure connection are perfectly achieved. Some tools have recently been developed that try and improve upon the user-interface problem faced here, and try to make the user

more aware of the fact that they may be using illegitimate certificates. These tools are discussed in Chapter 13.

The third method of deception is for the phisher to (a) ignore the fact that their illegitimate site does not provide security cues consistent with the legitimate site they are mimicking and (b) hope that the user does not notice. In other words, the phisher does nothing! Because a large number of lay users are ignorant of browsers' security cues and their meaning, this method is surprisingly effective.

Notice that in the PayPal phishing example previously covered in this chapter, this is the method of deception that was deployed. On the legitimate PayPal site you can achieve a secure HTTP connection by connection to `https://www.paypal.com` (note that the `https://` prefix indicates a secure connection, as opposed to traditional `http://` prefix that indicated an insecure connection). However, the actual link was `http://210.54.89.184/.PayPal/cgi-bin/webscrcmd_login.php` in the PayPal phishing example. This is an insecure connection and leads to a site controlled by the phisher.

1.4.6 The Catch

Once a phisher has collected confidential information from users, the information must be used for some sort of gain for the phisher; otherwise, why would the phisher go to the trouble of collecting the information in the first place? As alluded to earlier, there is generally a division of labor, and the criminal who makes use of collected information is not likely to be the same criminal who stole it in the first place. Nonetheless, a general understanding of how this information is used helps one understand the entire phishing ecology and thus can be useful in helping one when considering possible defenses.

In some cases, phishers target specific individuals and have a predetermined goal. For instance, in the case of industrial espionage a phisher might target a specific engineer associated with an aeronautics company, with the ultimate goal being to download schematics associated with a new wing design for a jumbo jet. With the specific catch in mind, the phisher's immediate goal might be to retrieve the engineer's authenticating data for the company's virtual private network (VPN). Once the authentication is retrieved, the phisher would use it to authenticate to the VPN and download the schematics of interest.

In the more generically targeted and common type of phishing scheme, where the goal is to retrieve the personal and financial information of as many victims as possible for the purposes of fraud, the value of and method by which the information collected by phishers is abused depends on the type and quality of the information. In general, the more personal and financial information that is retrieved for a given victim, the greater value it has. This is because more information permits criminals to abuse it with more degrees of freedom, increasing the chances that the criminals can successfully exploit the information. Having just the name and credit-card number of a victim is of increasingly little value to phishers, because online credit-card use is increasingly dependent on having a user's complete billing address, and credit-card security codes. A credit-card number with associated billing address and credit-card verification or security number can be used for the fraudulent purchase of goods or money transfers from online providers such as `Amazon.com`, eBay and Western Union. The risks to cashers in this form of fraud is that they have to have someone pick up the purchased goods or transferred money. For this reason, cashers often attempt to ship to either (a) international locations have no electronic fraud laws, or (b) countries that have such laws but in which they are unlikely to be enforced.

A check-card or banking-card number along with the issuing bank and associated PIN can be one of the most valuable collections of information that a phisher can get a hold of. The reason is that for some, but not all, banks the phishers are able to easily duplicate their bank's corresponding bank cards using easily acquired card-writers. Note that the difficulty in duplication of the card is not making the card itself, but the information and encoding on the card. Should the casher be able to duplicate the card appropriately, then they can use the card at an anonymous ATM and withdraw the maximal daily limit. The reason such collections of information are considered so valuable is that there is relatively little risk of detection or direct cost to the casher.

1.4.7 Take-Down and Related Technologies

Once a company or financial institution realizes that there is an active phishing attack on its customers or members, then there is an imperative to stop the phishing attack as quickly as possible, to protect as many customers as possible and minimize fraud. Generally this involves tracking down the computer that is hosting the hook website and having it removed, shut down, or made unaccessible. This is called *take-down*.

It is important to remember that typically the host of the phishing site is not a criminal, but rather simply another victim of the phisher whose computer has been hacked in order to host the phisher's website. Therefore, once it has been determined that a particular machine is involved in hosting a phishing site, it is normally sufficient to ask the administrator of the hosting computer to remove the offending website and ask them to practice better security procedures in the future to ensure that his or her computer cannot be used in any future attacks. Thus, once a phishing attack has been detected, it generally suffices to track down the computer and either (a) ask the offending host's Internet provider to block traffic or (b) ask for the host's administrator to remove the website in order to stop the attack. Of course, locating the appropriate computer and contacting the appropriate officials is not always an easy task, as the computer's domain name and/or IP address must be translated into sets of contact information for either the computer's network provider or administrators. Next, the corresponding officials need to be contacted and appropriately convinced that the machine is involved in an attack and that it must be taken down. This can often be a bureaucratic and/or legal mess, as people can be difficult to reach, and they may have legal or technical restraints that make disconnection or take down difficult, even if they intend to be fully cooperative in the take-down process.

It has been rumored that in some cases, such as when legitimate take-down was deemed to take too long or when the host's administrator and network provider are unreachable and/or uncooperative, certain service providers have actually launched denial of service attacks on the phishing sites, preventing potential victims from reaching the hook. Because of the questionable legality of such tactics, it is unlikely that any service provider will confirm such actions, but there are enough stories to suggest there is some grain of truth to them.

1.5 EVOLUTION OF PHISHING

The most immediate evolution in phishing attacks has been to polish and perfect the techniques that are currently in use. In the beginning, phishing lures were crudely worded emails often including many grammatical and spelling errors. Beyond that, the social engineering stories used to lure or convince people would be considered crude by todays standard, and the lures often lacked the professional graphic-designed appearance of modern phishing

lures. The result was that at the end of the day the early lures tended to be easy to distinguish from legitimate emails. Further, the phisher's websites were similarly crude, often looking nothing like the legitimate sites they were intended to mimic, containing obvious errors, and failing to have a polished look. In both the case of emails and lures, the phishers originally had few technical tools or tricks to overcome the apparent inconsistencies between their websites and the legitimate ones they were trying to mimic.

The division of labor in the phishing community has allowed for specialists to develop and master different portions of the attack. The result is the polished attacks we see today; however, the key structure behind the attacks of an email lure and hook website have remained largely unchanged during this evolution. Recently, we have seen the beginning of qualitatively and morphologically different forms of phishing attacks.

As previously hinted at, one qualitatively different form of phishing is known as *spear phishing* or *contextually aware phishing*. In such phishing, rather than sending the same lure to a large number of faceless victims, the phisher attempts to generate lures that are appropriate to the victim's context, or in extreme cases generate a shared context, so that the attack fits it. As an example of lures that are contextually appropriate, consider a victim who does all of their banking with the West Coast Bank and then receives a phishing lure claiming to be from the East Cost Bank. Since the bank does not correspond to the victim's context, she will easily distinguish it from a legitimate emailing and also possibly alert authorities at the East Cost Bank to the phishing attack. Now, consider an attack where the phisher makes an attack that mimics both the East Coast Bank and the West Coast Bank and before sending a lure to the victim determines which bank she uses. The likelihood that such an attack is successful is much higher than the original. The notion of contextually aware phishing attacks will be addressed in detail in Chapter 6.

A morphologically different form of phishing is given in the following case study. It gives an example of how a typical phishing attack can be evolved by replacing the typical email lure with an alternate lure based on commonly used search-engine-based web-shopping tools. It shows the diversity in phishing attacks we can expect to see in the future.

1.6 CASE STUDY: PHISHING ON FROOGLE

Filippo Menczer

Traditional phishing attacks to date have lured victims by spoofing email messages from banks and online vendors. A new, dangerous hook may originate from recent developments in e-commerce, namely the increase in confidence toward online shopping and the availability of easily accessible comparison shopping data.

Early generations of comparison shopping agents, or *shop-bots*, included systems such as *BargainFinder*, *PersonaLogic*, *ShopBot*, and later *MySimon*. They were aimed toward more efficient online markets [8, 3, 4]. From an implementation perspective, they either mined vendor sites for product information or asked vendors to pay a fee in exchange for having their offering listed (see [9] for a review). These methods implied various biases, but were difficult to manipulate by external third parties. The advent of shopping agents openly accessible via public APIs is making it possible —indeed easy— to manipulate product information for malicious purposes such as identity theft, by a form of phishing attack.

There are two aspects that make shop-bots vulnerable to being exploited for phishing. First, an attacker can lure shoppers into a phishing site by posting information about the fictitious sale of real products on a fake vendor site. Take, for example, Froogle, Google's "smart shopping" service (`froogle.google.com`). Anyone can submit product information for posting on Froogle by creating a free product feed through the Froogle Merchant program

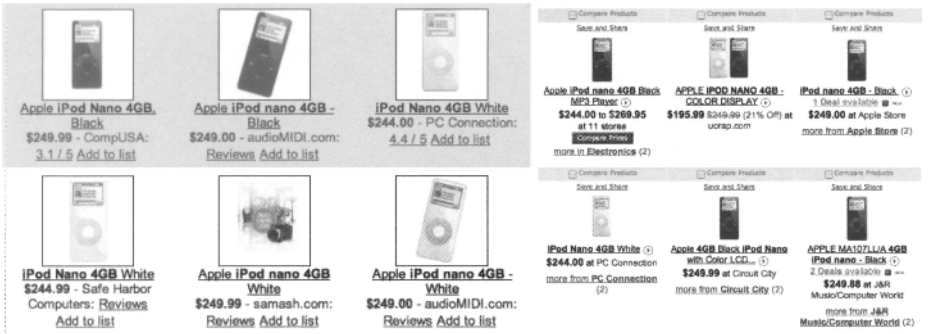


Figure 1.9 A few of the hits returned by Froogle (left) and Yahoo! Shopping (right) for the query "ipod nano 4gb."

(`www.google.com/froogle/merchants`). An attacker could advertise any popular item at an attractive site and thus direct a stream of visitors to the phishing site.

The second feature of open shop-bots (those with openly available APIs) that can be exploited by a phisher is the capability to determine the lowest price for a product. This is analogous to competitor analysis, whereby a firm collects valuable business information about its competitors through the web [10], but with the more malicious intent to advertise products at fictitious prices that are both attractive and credible. Take, for example, the Yahoo! Shopping site (`shopping.yahoo.com`). As with Froogle, users can interactively search through products and sort the hits by price (see Figure 1.9). But it is also possible to automate this process using the Yahoo! Shopping Web Service (`developer.yahoo.net/shopping`). A phisher's automated script can therefore calculate a credible low price in real time, both to keep the advertisement up-to-date with the lowest price on the market and to give potential victims a credible bid. Credibility is important because a price that is "too good to be true" may scare away prudent buyers, whereas one that is close to other vendors will not.

Once a victim is tricked into clicking on the link to the phishing site, the attacker can do several things. One possibility is to collect information about the victim, such as her browsing history (cf. Section 6.5), to be used in later attacks. Alternatively, the shopping deception can be continued by a fake e-commerce transaction to induce the buyer to complete the purchase, thus disclosing personal information such as a credit-card number. At this stage an even more insidious trick is to induce a victim to disclose his bank routing and account numbers. Many vendors offer an option to pay bills online via checking account as an alternative to a credit card. One way to induce the victim to select this option is to offer a further discount when such a payment method is selected. If the victim falls prey to this scheme, the phisher can wire money directly from the victim's bank account.

Demonstration To demonstrate the potential phishing exploitations of shop-bots, we have built a fictitious site called *Phroogle*.[2] The deception is illustrated in Figure 1.10. Users might be directed to this site by a shopping site (note, this is not actually done). In the demo the user can submit any query, and Phroogle fakes a search into its nonexistent database. Instead, the query is secretly sent to the Yahoo! Shopping Web Service and

[2] `http://newer.informatics.indiana.edu/cgi-b_n/phroogle/phroogle.cgi`

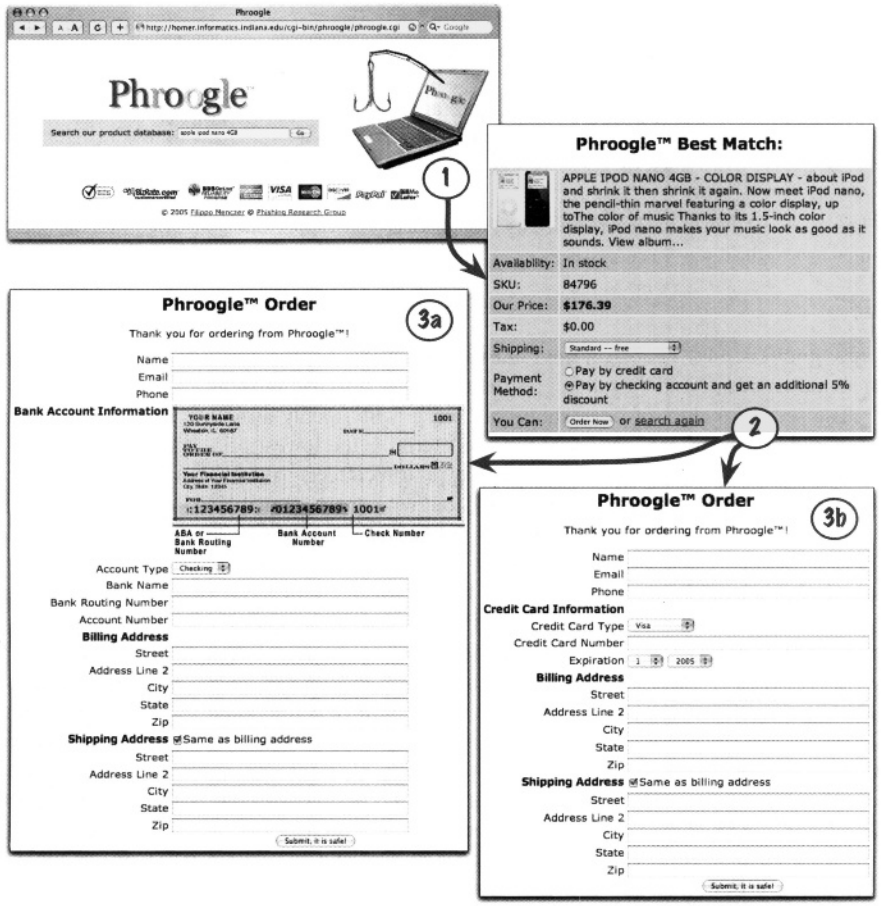


Figure 1.10 The Phroogle phishing deception: (1) The user submits the query "ipod nano 4gb." Note the fake logos to give the victim a false sense of trust. (2) Phroogle forwards the query to a shop-bot service and presents to the user a real product picture and description returned by the bot, along with a fake price lower than those from real merchants. A further discount is offered for paying by checking account rather than credit card. (3) Once the victim selects a payment option and clicks on the order button, personal information has to be supplied—this would be the actual theft perpetrated by the phisher.

the user receives a hit in real time. The hit looks real, as it contains an actual product image and description returned by Yahoo. But the price is 10% below the lowest price found (cf. second Yahoo hit in Figure 1.9), so that it looks like a very good, yet credible, offer. Phroogle offers another 5% off the listed price if the user pays with bank account information, to induce the victim to disclose that information. If the user instead selects to pay by credit card, the phisher would steal the credit card number and simultaneously obtain information on the purchaser's email and billing address; this may be useful in other scams as well.

Phroogle is meant as an innocuous demo rather than an actual phishing attack. We give users a big hint about its phishing nature using an obvious icon and a link to our phishing research group site. We do not collect any personal information. We simply count the number of people who complete each stage of the transaction and offer summary statistics to those users who complete the transaction. These numbers are not significant because the visitors to the site are aware that this is a phishing attempt. Nevertheless, Phroogle demonstrates that a phisher could easily exploit shopping agents to set up an effective phishing attack.

Malware and Its Relation to Phishing

Another direction in which phishing attacks are currently evolving is that the phishing attack is being combined with different types of malware such as Trojan horses or key-logging software. With appropriate Trojans and key-logging software installed on a victim's machine, phishers can wait until a user visits a legitimate website, such as their online banking site or a favorite ecommerce site, and then steal the username and password used to access the *legitimate site*. Additionally, at this point the malware can steal any financial information that might be entered *or displayed on the screen*.[4] We note that such attacks could be devastating, as they could potentially give far more information to the phisher than they currently have access to. Consider the possibilities for a phisher if he had access to the information that is displayed on an online banks website related to a user's credit card. Such a site normally contains not only the credit-card number, but information related to prior purchases and the credit limit of the card. Such information could be used by the phisher to make the purchases on the card look much less suspicious to the credit-card companies' fraud detection systems, and thus to maximize the possible fraud.

Clearly, there can be some damning phishing attacks that can be conceived of with the complementary use of malware. However, the editors of this book have chosen to try and limit, if not ignore, the discussion of such attacks. This decision was not made because such attacks should be considered harmless or not worthy of discussion, far from it! The reason for the minimization of such coverage is that the topic of defenses against malware is worthy of its own book, and it deals far more intimately with issues related to secure operating systems, trusted computing platforms, and bug-free and secure coding practices. Therefore, while no clear line can be drawn to divide traditional phishing from phishing with malware, we have chosen not to emphasize this topic. Additionally, even if malware can be eliminated, it is the editors' belief that many of the attacks and discussions in this book will still be applicable and thus must be addressed. That being said, a large portion of Chapter 4 is devoted to discussing the basic ideas behind different types of malware.

[4]With the use of malware, phishers can actually capture any information displayed on the screen, so in these cases the user has to be cautious of not only what is entered into the website, but also what is displayed on her screen.

without getting in to too many specifics, so readers can have a general idea of the dangers such softwares represent.

1.7 PROTECTING USERS FROM PHISHING

Given the growing problem of phishing, it is clear that the problem needs to be addressed, and defenses need to be deployed to protect users. And, while undoubtedly there is much that can be done to protect users and online service providers from such attacks, there is unlikely to be any silver bullet that can completely prevent them. Part of the reason is that phishing relies on social engineering, and people can often be convinced to do things that are completely detrimental to their well-being, if asked to them in an appropriate manner. Take, for example, the following hoax email that was going around the Internet several years back:

> I found the little bear in my machine because of that I am sending this message in order for you to find it in your machine. The procedure is very simple.
>
> The objective of this e-mail is to warn all Hotmail users about a new virus that is spreading by MSN Messenger. The name of this virus is jdbgmgr.exe and it is sent automatically by the Messenger and by the address book too. The virus is not detected by McAfee or Norton and it stays quiet for 14 days before damaging the system.
>
> The virus can be cleaned before it deletes the files from your system. In order to eliminate it, it is just necessary to do the following steps:
>
> 1. Go to Start, click "Search"
> 2. In the "Files or Folders option" write the name jdbgmgr.exe
> 3. Be sure that you are searching in the drive "C"
> 4. Click "find now"
> 5. If the virus is there (it has a little bear-like icon with the name of jdbgmgr.exe DO NOT OPEN IT FOR ANY REASON
> 6. Right click and delete it (it will go to the Recycle bin)
> 7. Go to the recycle bin and delete it or empty the recycle bin.
>
> IF YOU FIND THE VIRUS IN ALL OF YOUR SYSTEMS SEND THIS MESSAGE TO ALL OF YOUR CONTACTS LOCATED IN YOUR ADDRESS BOOK BEFORE IT CAN CAUSE ANY DAMAGE.

The file referred to in this hoax email is actually a legitimate file related to Java on the Windows operating systems, and it is installed by default on Windows machines. Therefore, everyone receiving the email who believed that the email was convincing found the file and probably deleted it. Since the file is not considered an essential operating system file, the worst-case result in such a situation is that some Java applets may not have executed properly, and the `jdbgmgr.exe` file would need to be reinstalled to regain the applets' functionality. However, it is simple to abstract this attack and imagine a similar message arriving by email telling users to delete software designed to counter certain phishing attacks, followed several days later by the lure to such a phishing attack. In general, phishers can use social engineering in order to ask users to use their systems in ways that were never intended by system designers. It will be hard, if not impossible, for security engineers to design against this, while still making it possible for users to easily accomplish the tasks they expect to with their computers.

In the coming chapters the reader will be exposed to many different phishing attacks and associated techniques and tricks used in these attacks. Similarly, the reader will be

exposed to countermeasures that researchers in the field are developing. Many of the current countermeasures are in their infancy, but when developed will probably do much to minimize the risk of the current forms of phishing attacks. The expectation among researchers is that countermeasures that simply solve today's phishing attacks will be quickly bypassed as phishers evolve their attacks. Thus, when designing countermeasures, it is essential that researchers try to consider not only current attacks, but the weaknesses in the infrastructure that such attacks take advantage of, and how to protect these weak points. This will ensure that simple evolutions of the phishing attacks do not overwhelm the proposed countermeasures. Therefore, contributors have endeavored to discuss current attacks not only in their current embodiments, but in their expected or natural evolutionary forms. Similarly, countermeasures are discussed that might be created and used to stop these future phishing variants.

REFERENCES

1. Christopher Abad. The economy of phishing: A survey of the operations of the phishing market. *First Monday*, 10(9), 2005.
2. Alice Dragoon. Foiling phishing. *CSO Online*, October 2004.
3. A. R. Greenwald and J. O. Kephart. Shoppers and pricebots. In *Proc. 16th Intl. Joint Conference on Artificial Intelligence*, pages 506–511, 1999.
4. J. O. Kephart and A. R. Greenwald. Shopbot economics. *Autonomous Agents and Multi-Agent Systems*, 5(3):255–287, 2002.
5. Karen Krebsbach. Goin' phishin. *Bank Technology News*, April 2004.
6. Avivah Litan. Phishing victims likely will suffer identity theft fraud. *Gartner Group*, (M-22-8474), 2004.
7. Avivah Litan. Increased phishing and online attacks cause dip in consumer confidence. *Gartner Group*, (G00129146), June 2005.
8. P. Maes, R. H. Guttman, and A. Moukas. Agents that buy and sell. *Communications of the ACM*, 42(3):81–91, 1999.
9. U. Menezer, A. Monge, and W.N. Street. Adaptive assistants for customized e-shopping. *IEEE Intelligent Systems*, 17(6):12–19, 2002.
10. Y. P. Sheng Jr and P. P. Mykytyn and C.R. Litecky. Competitor analysis and its defenses in the e-marketplace. *Communications of the ACM*, 48(8):107–112, 2005.
11. Elizabeth Robertson. Phishing victims likely will suffer identity theft fraud. *Tower Group*, December 2004.
12. TRUSTe and the Ponemon Institute. *Press Release*, September 2004.