

# The Human Factor in Phishing

Markus Jakobsson  
School of Informatics  
Indiana University at Bloomington  
markus@indiana.edu

## Abstract

We discuss the importance of understanding psychological aspects of phishing, and review some recent findings. Given these findings, we critique some commonly used security practices and suggest and review alternatives, including educational approaches. We suggest a few techniques that can be used to assess and remedy threats remotely, without requiring any user involvement. We conclude by discussing some approaches to anticipate the next wave of threats, based both on psychological and technical insights.

## 1 What Will Consumers Believe?

There are several reasons why it is important to understand what consumers will find believable. First of all, it is crucial for service providers to know their vulnerabilities (and those of their clients) in order to assess their exposure to risks and the associated liabilities. Second, recognizing what the vulnerabilities are translates into knowing from where the attacks are likely to come; this allows for suitable technical security measures to be deployed to detect and protect against attacks of concern. It also allows for a proactive approach in which the expected vulnerabilities are minimized by the selection and deployment of appropriate email and web templates, and the use of appropriate manners of interaction. Finally, there are reasons for why understanding users is important that are not directly related to security: Knowing what consumers will believe—and will not believe—means a better ability to reach the consumers with information they do not expect, whether for reasons of advertising products or communicating alerts. Namely, given the mimicry techniques used by phishers, there is a risk that consumers incorrectly classify legitimate messages as attempts to attack them. Being aware of potential pitfalls may guide decisions that facilitate communication.

While technically knowledgeable, specialists often make the mistake of believing that security measures that succeed in protecting *them* are sufficient to protect average consumers. For example, it was for a long time commonly held among security practitioners that the widespread deployment of SSL would eliminate phishing once consumers become aware of the risks and nature of phishing attacks. This, very clearly, has not been the case, as supported both by real-life observations and by experiments [48]. This can be ascribed to a lack of attention to security among typical users [47, 35], but also to inconsistent or inappropriate security education [12]—whether implicit or not. An example of a common procedure that indirectly educates user is the case of lock symbols. Many financial institutions place a lock symbol in the *content portion* of the login page to indicate that a secure connection *will be* established as the user submits his credentials. This is to benefit from the fact that users have been educated to equate an SSL lock with a higher level of security. However, attackers may *also* place lock icons in the content of the page, whether they intend to establish an SSL connection or not. Therefore, the use of the lock

symbol in the content part of the page dilutes the importance of the true SSL lock symbol. To many users, it is not clear exactly where the lock symbol needs to be placed to signal security.

Educating consumers that the lock must appear in the address bar or the chrome of the webpage and raising their awareness of phishing does not eradicate the problem: It has been shown that the typical computer user is unable of distinguishing a valid certificate from one that is invalid or self-signed (see, e.g., [46, 40]), and that he may not even detect the absence of indicators that a connection is SSL secured. The latter statement is supported by a recent study [25] that showed that while users often detect the *presence* of *incorrect* information, they almost never detect the *absence* of *correct* information.

## Experimental Assessments of Consumer Psychology

Fogg et al. [14, 15] conducted a study with over 2500 subjects and investigated how different elements of websites affect people’s perception of credibility and laid down guidelines for the credible perception of websites. Dhamija, Tygar and Hearst [7] recently studied how computer users fall victims to phishing attacks based on a lack of understanding of how computer systems work, due to a lack of attention, and because of visual deception practiced by the phishers. Their research involved finding out what indicators of security people are looking for while judging a website as fake or authentic, and their study provides a quantitative measure of how these aspects manifest themselves in people’s susceptibility to attacks. In their user survey, they noted that 23% of subjects completely overlooked browser based security clues such as the address bar, the status bar and the SSL lock icon, and that 40% of subjects made the wrong security decision. We report on extensions of their study in which the thought process leading to security decisions is also analyzed, and the importance of design issues is quantified.

To understand what typical computer users react to, and why, we performed two in-lab user studies [28, 44]. Subjects in both of these studies were shown emails and webpages, and were asked to rate these in terms of their likely authenticity. Subjects were shown both legitimate stimuli and stimuli that corresponded to attempts to deceive, and asked to rate these on a scale from 1 (very phishy) to 5 (very likely to be legitimate).

It is clear that studies of these kinds are bound to introduce a bias, as subjects know that they are being tested on their ability to detect phishing. Therefore, their awareness is heightened. However, we argue that such studies are still useful to obtain insights into *relative* appearances of security: One can compare the reactions to different stimuli in studies of this kind. We believe that in order to determine what makes emails and webpages believable (as opposed to *to what extent* they are believed), in-lab user studies are probably better suited than naturalistic experiments (e.g., [13, 21, 25]). This is due to the fact that naturalistic experiments typically can not test a sequence of stimuli for each subject, which increases the required sample size dramatically. Therefore, in spite of the fact that naturalistic experiments can offer bias-free results, in-lab experiments may still be preferable. The exit interview provided an opportunity to determine what approaches might have been more suitable, based on the answers given by subjects.

In a first study [28], a qualitative approach was taken in which subjects *spoke their thoughts aloud* as they rated stimuli. In an exit interview for the study, we also asked subjects what would have made them react differently to stimuli they found noteworthy.

Guided by the ratings given to different stimuli, we reviewed the recordings for clues of why stimuli were interpreted as they were. We also attempted to determine what caused subjects to make up their minds about the trustworthiness of stimuli. Typically, this decision was made right before a classification was performed. We took note of pivotal observations appearing to inform the decisions of subjects. Having collected a large number of pivotal observations, we then interpreted these in the context of the quantitative ratings to find a likely implication. The implications, in turn, correspond to conclusions about how subjects make decisions of trust.

Many of these conclusions support already held beliefs, whereas some highlight aspects that are not common knowledge, and some even contradict commonly held beliefs.

In a second study [44], a quantitative approach was taken instead, in which two large groups of subjects—totaling close to 400 subjects—were shown sequences of stimuli such that the stimuli shown to each group were near identical. The differences in the ratings of near-identical stimuli were used to assess the importance and impact of the differences between such related stimuli. This way, we could simply test the impact of minor design differences, whether these are actually in use or simply held a promise to be of relevance. Such differences related to the design and presentation of material, the inclusion of various types of trust indicators, and various personalization aspects.

The conclusions from the two studies are presented below. When reading the conclusions, it is important to keep in mind that these were made in the context of subjects who knew that they were being evaluated on their abilities to detect phishing attempts; therefore, they describe the *abilities* of the subjects rather than the *habits* of the subjects. This means that some of these observations may not hold in a real-life setting, unless the users in question are on the lookout for potential attempts to attack them.

1. **Spelling and design matter.** The number one aspect that subjects consider is the design and spelling of messages. Several phishing emails were dismissed based on spelling alone. Subjects rarely paused to notice third party endorsements – let alone alter their judgment – in the presence of a gross grammatical error. Many subjects were suspicious of emails that were not signed by a person (Jim Smith) but instead by a position only (e.g., “Account manager, Paypal”). Similarly, subjects criticized email messages that instructed them not to reply. Some legitimate providers (such as Keybank) were given a low rating due to “unprofessional design.” In the case of Keybank, subjects cited the absence of the institutional name on the login-page, along with the fact that the fields for user name and password were of different length. The presence of copyright information and legal disclaimers, typically at the bottom of the stimulus in small print, enhanced trust for many subjects quite dramatically. Subjects argued that phishers do not need legal disclaimers, and do not care about copyright infringement. Therefore, phishers are not likely to include such texts in messages, the argument was.
2. **People look at URLs.** It was found that subjects looked carefully at URLs of webpages, and on the URLs obtained by mouse-over in emails. Subjects were good at detecting IP addresses as being illegitimate URLs, but were not highly suspicious of inauthentic URLs that were well-formed, such as `www.chase-alerts.com`. On the other hand, subjects were good at detecting syntactically peculiar addresses, such as `www-chase.com`. (Whereas this is a well-formed URL, most subjects did not know this, and treated it much like a spelling error.)

A recent naturalistic study [25] found that the yield of a simulated attack more or less doubled when the URL shown at mouse-over and in the address bar was believable; this supports that URL viewing *is* used for real security decisions, in clear contrast to what is often believed. If we compare the results of [25] (which corresponds to a naturalistic study) to those in [28, 44] (which are in-lab studies), then we see evidence that the reliance of mouse-over is greater in the in-lab experiment. This supports that the security *abilities* of typical users differ from their security *habits*, and we confirm that subjects behave differently when they *know* that they are participating in an experiment.

It was found that subjects favor short URLs. Tsow et al. [44] showed that a page with URL `https://www.accountonline.com/View?DocId=Index&siteId=AC&langID=EN` was considered significantly *less* trustworthy (with  $p < 0.004$ ) than a page whose URL was `http://www.attuniversalcard.com`. Here, the *contents* of these two pages were identical, and the first page was actually SSL protected, but was still given a lower rating.

3. **Too much emphasis on security can backfire.** Some stimuli were criticized for their overwrought concerns about online security. An example of this is the IU Credit Union website shown in Figure 1; this is a legitimate website, but with a very strong attention to phishing. Subjects did not like that this website said “phishing attack in progress” in three different locations. Some commented that “phishing” is too obscure a term for a financial institution to use in their communications – the phrase “identity theft” was offered as a plausible substitute. In Tsow et al. [44], it was established that if the focus on security was downplayed, then there was a significant increase in trust ( $p < 0.022$ ).



Figure 1: The top part of the actual webpage of Indiana University Credit Union. No subjects noticed that the URL was incorrect—[www.IUCU.com](http://www.IUCU.com) as opposed to the real [www.IUCU.org](http://www.IUCU.org)—but many felt that the site probably was not legitimate, citing the unusual emphasis on security as their reason. Lower down on the same page, phishing was mentioned again.

4. **Third party endorsements depend on brand recognition.** Several stimuli included a range of third party endorsements, from well established brands like Verisign to made-up endorsements like *Safe Site*. We found that endorsements from Verisign were taken with the most gravity. Almost every subject mentioned Verisign by name as a positive factor in their trust evaluation. BBBOnLine and TRUST-e endorsements had no significant effect. Made-up endorsements evoked consistent criticism.

Some subjects noticed third party endorsements on stimuli they clearly believed to be phishing, and deduced that the graphics could be rendered on any page. One subject observed “Probably now that I see all these [stimuli], I should not believe in Verisign,” but later dismissed a web page because “it’s not Verisign protected, but it says something which I’ve never seen, ‘TRUST-e’. I don’t know, so probably I wouldn’t go in this account.” No other third-party endorsement was mentioned by name as a prerequisite for trust.

5. **People judge relevance before authenticity.** Subjects often decided whether a stimulus was legitimate or not based on the content, as opposed to the signs of authenticity. In particular, any stimuli that offered a monetary reward was considered phishy, independently of whether it was authentic or not. Likewise, emails that requested passwords upfront were considered phishy, whereas emails that only appeared to contain information were considered safe. The latter is a problem, as users could be drawn to a site by an email that appears to be for information only, and once at the site, asked for credentials. Having already made a trust decision (for the email), we believe that the user is now less likely to be critical of the webpage material.

We note that the observation that people judge relevance before authenticity may pose a problem to companies that rely on surveys or on advertising that is likely to be considered phishy by recipients.

6. **Personalization creates trust.** A high degree of personalization increases the trustworthiness of stimuli, whether email or webpages. Thus, the more personal information is present, the more likely did the subject find that the stimulus was authentic. This suggests that data mining could be a troublesome new aspect of phishing. One subject said that the presentation of ZIP code and mother’s maiden name would enhance trust in an email message. Yet, this data could be gathered by an attacker using geolocation software (such as [18]) and publicly available databases—see [17] for a recent study on the security of mother’s maiden names.

Many financial service providers attempt to authenticate themselves to their clients by including information of the last few digits of the account number of the client. However, no legitimate service providers use the first few digits of an account number as an authenticator, as these digits typically are identical for large numbers of their clients, and so, can be anticipated by an attacker. Subjects did not realize this, and some found it comforting with an email stating it was intended for a user whose account starts with 4546. Many subjects insisted that presence of the last four digits is more trustworthy, but did not penalize a message for using the first four digits. Some commented that they did not like to see the prefix in isolation, but preferred it to be formatted with the others starred out, e.g., 4546-\*\*\*\*-\*\*\*\*-\*\*\*\*.

7. **Emails are very phishy, web pages a bit, phone calls are not.** Overall, email stimuli were considered more phishy than web stimuli to participants in the study. Many subjects said that following links from email was a risky activity, and that they consciously avoid the practice. Since very few admit to following links given in phishy emails, it could be assumed that their exposure to phishy webpages is inherently more limited. Many participants said that they would try to independently verify email contents by calling the institution directly. Few participants specified how they would obtain the correct phone number and therefore could expose themselves to fraudulent customer service numbers. We note that most systems prompt users to dial in their account number and zip code prior to speaking with a representative, which trains consumers to be ready to give out identifying information and credentials when initiating a phone call to a financial institution. Several participants also said that email is an inappropriate alert medium for urgent matters, such as password changes and account lock-outs, and expected a phone call from the institution. A strategy using automated phone messages may increase an attack’s potency. For example, a voicemail alerting potential victims to the imminent receipt of an email may improve the consecutive email message’s apparent legitimacy.

8. **Padlock icons have limited direct effects.** Large padlock graphics were effective at drawing attention to specific portions of the stimulus. By themselves, they did not cause any subject to express an improvement in trust. Small padlock icons in the content body were never commented on by subjects. Their ineffectiveness was supported by the nearly identical rating distributions of two Chase web pages that differ only by the presence of the SSL-post padlock icon in the login area. One of the pages was the real Chase login page, and the other one was a slight modification of the same in which the padlock graphics was removed.

The SSL padlock at the bottom of a browser frame enhanced trust in many subjects, however two subjects lost trust when mouse-over revealed a made up certification authority, *Trust Inc.*

Most users were confused by the presence of a favicon<sup>1</sup> padlock in the browser’s address bar. We were surprised by this result because we hypothesized that the address bar contents would be more trusted since web servers have limited control over its appearance.

9. **Independent channels create trust.** If a stimulus suggested that the subject could call to verify the authenticity of the email/webpage, then the very existence of this possibility strengthened the trust the subjects had in this stimuli. Subjects stated that they would not call the number to verify the authenticity, but “*someone else would*”.
10. **People recognize common attacks.** Subjects were better at detecting potential attacks that were of commonly occurring types than structurally similar attacks using different language. This suggests that while people may become better at recognizing phishing, they are no less vulnerable to attacks that use different deceit approaches. Therefore, as soon as a new psychological twist is developed, there is reason to believe that it will become successful.

## Educational Efforts

While educational efforts designed to reach typical computer users can help change the way consumers react to phishing, there are also inherent limitations in what can efficiently be communicated, given the complexity of the problem and the relative lack of interest in active involvement on behalf of typical users. While “phishing IQ tests” [31] have been proposed as a way to measure the efficacy of phishing education, recent studies [2] suggest that they fail to measure the ability to recognize threats. However, they do show that exposure to traditional phishing education raises the level of concern among consumers—at least in the short run.

There is an abundance of efforts attempting to educate consumers of the risk of phishing, spanning the spectrum between popular articles (such as [38]) and books (e.g., [1]). Most of these efforts are traditional, and are believed among security researchers to have rather limited impact. There are some recent efforts to develop non-traditional methods to educating consumers, with promising efforts based on computer games by Kumaraguru et al. [30], and an effort using a comic-book format by Srikwan and Jakobsson [42]; an example of the latter effort is shown in Figures 2 and 3. Common to both of these efforts is the attempt to relate the educational message to the audience in a manner that encourages the audience to learn more. The comic approach relies heavily on analogies and on reviewing common threats in a way that the reader can relate to.

---

<sup>1</sup>The favicon is the small logo used in the address bar by many organizations. This logo can be set to a lock, to make it appear that the page is SSL secured.

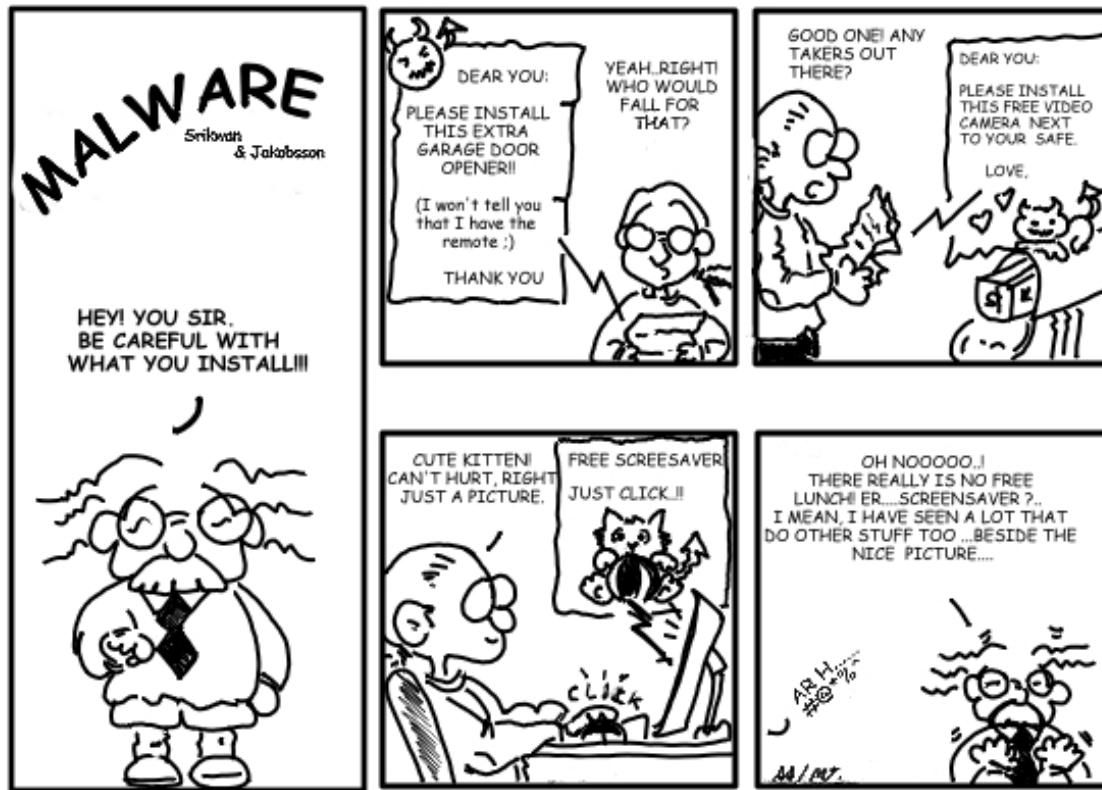


Figure 2: A recent and non-traditional effort to explain security threats to consumers [42]. The strip is part of a larger sequence of strips on malware, and is continued in Figure 3. In order to explain online threats, the comic relies heavily on real-world analogies. In the first and second panels, one of the main characters, the “smart guy”, is confronted with real-world versions of spyware. Just like most consumers, he can make appropriate security decisions in the real world, but finds it difficult to translate real-world wisdom to online threats. In the third panel, he is tricked to install spyware in the guise of a screen saver. The good professor, who is used to provide commentary, laments the decision of the smart guy in panel four.

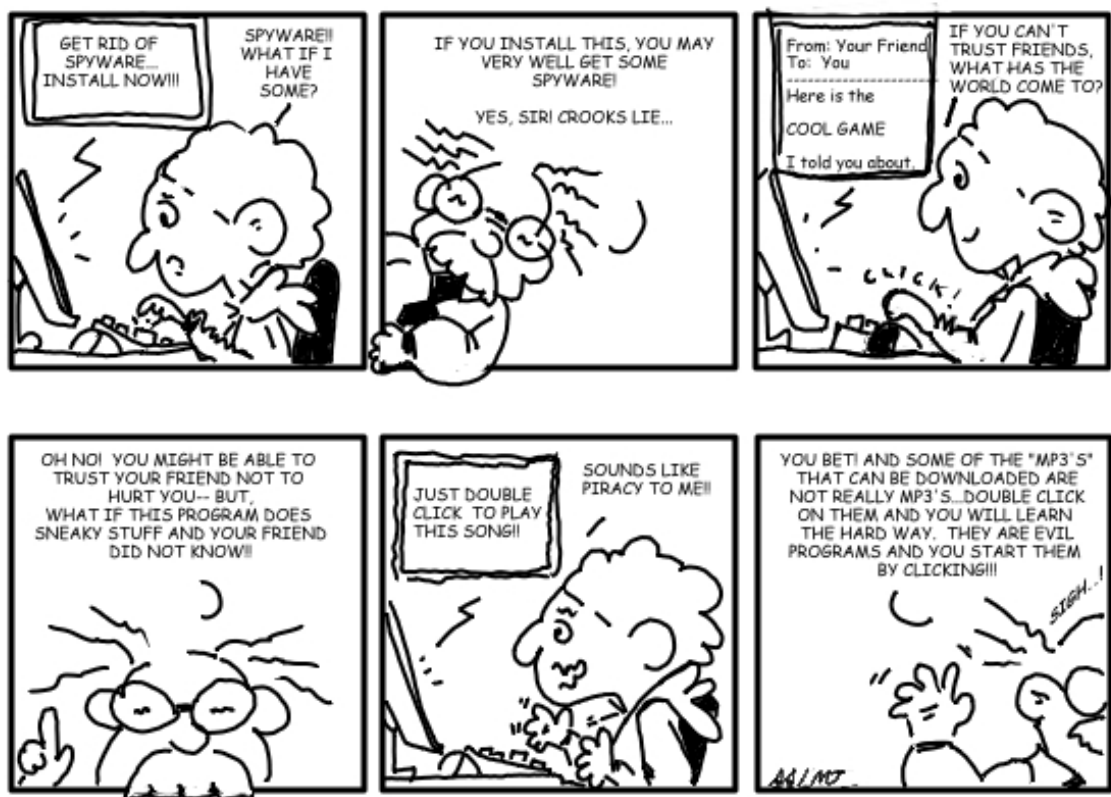


Figure 3: A recent and non-traditional effort to explain security threats to consumers [42]. The strip is part of a larger sequence of strips on malware, and is a continuation on the story started in Figure 2. In the first panel, one of the main characters, the “worried guy”, is exposed to the risk of spyware. In the second panel, the good professor comments on the situation. In the third panel, the worried guy receives an email from his best friend, suggesting he installs a game; this is commented on by the professor in the fourth panel. In the fifth and sixth panel, the strip treats the risk for malware that is posed by piracy. The worried guy is about to click on a purported MP3 file, not realizing that this might instead be an arbitrary executable with an icon that is designed to look like an MP3 file. While this strip emphasizes risks and what *not* to do, other strips instead give suggestions of what corresponds to safe behavior, especially in the context of potential attacks.

## 2 Keeping Up with Psychological Threats

In order to defend against new types of deceit, we believe it is best to anticipate the threats, whether by studying existing attacks on other brands, or by employing people in charge of “being” the attacker—i.e., trying to think a step ahead all the time. Once the likely threats have been established, these may be countered by designing user interfaces in a way that promotes security; by pre-emptively registering domains that could benefit attackers; and by other approaches that limit the freedom to attackers.

Here we present some strategies we believe may be useful to mitigate deceit-based attacks; this is in no way an exhaustive list of defenses, but rather, is intended to illustrate some potentially helpful approaches.

### Mitigating side-channel threats

Our experimental data suggests that the use of side-channels, such as regular telephony, may increase the yield of attacks. Assume for a moment that an attacker can make an educated guess about the likely banking affiliation of a given user (see section 3 or [22] for a description of how this might be done), or simply is lucky to pick the right one. Consider then an attack in which each potential victim receives a sequence of emails appearing to come from his or her bank, each email being a notification of some fictitious transaction, and in which the phone number to call in case of questions has been changed to one that the attacker controls. Most banking clients are used to interacting with an automated system when calling their bank, and, according to informal surveys, most people would feel that since they initiated the call (and “know” whom they are calling), it is reasonable for them to enter credentials to get to speak to a person. While this in essence is structurally the very same attack as users are becoming weary of when it is perpetrated on the Internet alone, it has two new twists: First, the attack emails (the “lures”) do not demand that the user acts, but rather, are purely for his or her information. Second, instead of relying on a recipient visiting a website, the attack is based on getting victims to place a phone call. Apart from the fact that people appear more willing to give out information when they feel they initiated the transaction, this hypothetical attack benefits from the fact that take-down—while becoming reasonably fast for websites—is not likely to be very fast at all for phone numbers. How is an attack like this best avoided? We believe that the answer to this question is to avoid training clients to call the number indicated in the email if they have questions, but rather, always ask clients to use the number on the back of their credit card, ATM card, on a recent statement, or one that they look up themselves. Doing so will not *stop* an attack of this type, but will at least avoid *helping* the attacker.

### An eye on mutual authentication

Clients of financial institutions are used to having to authenticate themselves in order to obtain service. As a result of the rising tide of phishing attacks, they are also becoming used to financial institutions providing some authenticating information *to them*. This is commonly done by stating the last few digits of an account, but there are efforts to deploy methods that offer better security guarantees, e.g., by presenting user-specific images to clients after having verified their IP addresses, cookies, and other partially identifying information. Whereas both approaches offer some security benefits, we argue that neither is bullet-proof.

As described in section 1, few subjects distinguished between authentication based on the *last* few digits of an account number and the *first* few. The latter, clearly, offer a less meaningful manner of authentication, as large batches of accounts (if not all) have the same first few digits for a given issuer or financial institution. Since the last *first* digits do not provide a meaningful way of authentication and typical computer users do not notice one approach being switched for the other, then authentication using the *last* four digits arguably invites abuse by training

clients to accept a form of authentication that is easily abused. Similarly, stating the email address to which the account is registered clearly does not provide any degree of authentication, but offers the *perception* of authentication to clients, which makes it dangerous to do—whether it is intended for authentication or not.

Another approach is to use a user-specific image to authenticate to the client, as is done by Bank of America’s SiteKey [39]. This is a promising approach, but preliminary experiments support that attackers could still use deceptive approaches in which the images are switched. Namely, subjects found the following message believable: *“Due to the Americans with Disability Act (ADA), we are replacing all images with high-contrast images. Your new image is presented below. To acknowledge this change, please log in at [www.bankofamerica.image-update.com](http://www.bankofamerica.image-update.com) at your earliest possible convenience, and check the box indicating that you agree to using the displayed image. Before logging in, please verify that the image is the same as what is presented below.”*

While attacks that involve the deceptive switching of authenticators are possible when bank-chosen authenticators are used, there are also distinct drawbacks associated with allowing users to provide the authenticators during a setup phase. Apart from the increased burden carried by clients and the associated support calls, there is also the problem of poorly chosen authenticators. For example, if users select the images later used to authenticate the bank to them, many may choose a picture of their favorite musician, which significantly reduces the overall entropy of the set of authenticators. One way to approach the problem is to use information that means something to the client, but which the client does not have to explicitly select and provide. We suggest that an alternative approach would be to use the billing address of the client as a means for authentication. The benefits of this approach is that it is very difficult for attackers to associate physical addresses and email addresses on a large scale. Given potential privacy concerns of consumers, however, one could use *part* of the address only, instead of the full address. Still, this is not a panacea: An attacker may simply send spoofed emails in which the address-based authentication is absent (see [25]). Clearly, the authentication has to be made prominent enough that its absence is noticed. Also, one must remember that attackers may attempt a two-round attack, the first phase of which involves tricking the victim to give out his mailing address. Still, this increases the burden on the attacker in comparison to one-round attacks.

## Copycat defense

There are many ways to keep up with the threats and anticipate the next wave of attacks. First of all, history shows that after a new type of attack has been successfully launched against one brand, it is soon launched on other brands as well. For example, in the spring of 2006, an attack targeting Chase customers involving a request to fill a survey became common. In the late spring the same year, similar attacks targeting WaMu started to appear. There is evidence [34] suggesting that the second round of attacks were not performed by the same attacker, but by a copycat. One such piece of evidence was that whereas the attacks on Chase involved sites hosted at sites with deceptively named cousin-names, such as **chase-rewards.com**, the webpages involved in the WaMu attack instead had URLs that were based on plain IP addresses. This is known [25] to be substantially less effective, but requires less sophistication on behalf of the attacker. While the reason could have been that all suitable cousin-domains had been pre-emptively registered by WaMu, this is not the case: We registered **wamu-rewards.com** at the onset of the attacks (and have, half a year later, still not been asked to transfer the domain.) One way of keeping up with the threats—and lower the expected yield of copycat attacks—would be for brands to pre-emptively register all suitable cousin-name domains relating to the psychological twist of a new attack. It is commonly argued that this is infeasible, as there is a near-endless number of possible domains; however, there is a fairly limited number of domain

names that *closely match* the structure and concept of a given attack. In the case of the survey-based attack, each brand not yet affected by this particular attack could have monitored all the domains used by the attacker on Chase, substituted the term “chase” with their own name, and registered the resulting domain. This requires no particular effort, given that most financial institutions run honeypots in the first place. Requiring only slightly more effort, each brand could set out to also register all cousin names that are related to the psychological twist, as determined by people observing the first attack.

## Defense against feature-based attacks

A second way to benefit from pre-emptive registration of cousin-name domains is to register domains that match existing or potential services or features that are or have been offered by the brand or its competitors. For example, Chase offers an alert service in which notifications are sent to clients who opt in to this service. The clients are sent with an apparent sender of `Chase@alerts.Chase.com`. We registered `alerts-chase.com` and `chase-alerts.com`; the domains were later requested by and transferred to Chase. These domains were registered since, presumably, they could otherwise have been used by an attacker who sent out legitimate-looking notifications corresponding to fictional transactions, adding a line such as “*For more details on this transaction, log in to see your alerts at [www.alerts-chase.com](http://www.alerts-chase.com)*”. While this format would be slightly different than the actual notification (in which no link is given), it is believed that very few recipients would notice the difference, or care about it if they did notice.

While it is more difficult to anticipate attacks based on legitimate advertising campaigns, this is still meaningful. For example, an attacker might register `switch-to-citi.com` and use this in an apparent effort to have the recipient open an account with Citibank. An example of how such an email lure might look is shown in Figure 4. This type of attack has the side benefit (as far as the attacker is concerned) that he does not have to target people who *are* clients of Citibank, but rather, people who are *not*. This, arguably, makes his email believable to a larger group of potential victims. The attacker could use this either in concert with an existing advertisement campaign, or independently of the same. He would be able to offer very nice enticements for people who agree to make the switch (which is made simply by “transferring” a small amount from an existing bank account). This highlights why phishers often have higher click-through rates than legitimate providers of advertisements: Fraudsters can offer much nicer enticements than legitimate service providers, as they are not tied to their word.

## Minding acquisitions

It is known that the time just around the acquisition of a competitor increases the vulnerability to attacks, as do any other big changes that can be observed by consumers. As the acquisition is announced, it makes sense to register domains that might otherwise be taken advantage of, such as domains that are the concatenations of the names of two financial institutions, and slight variations of these, e.g., `bankone-becomes-chase.com`. It also makes sense to protect domain names that correspond to functional changes that become apparent to consumers. For example, assume that the financial institution using six-digit PINs is acquired by a financial institution using four-digit PINs. It is imaginable that clients of either institution would be requested to change their PIN. We preemptively registered `PIN-update.com` and related domains, anticipating that these could be used in such an attack, potentially using the name of the attacked brand as a sub-domain.

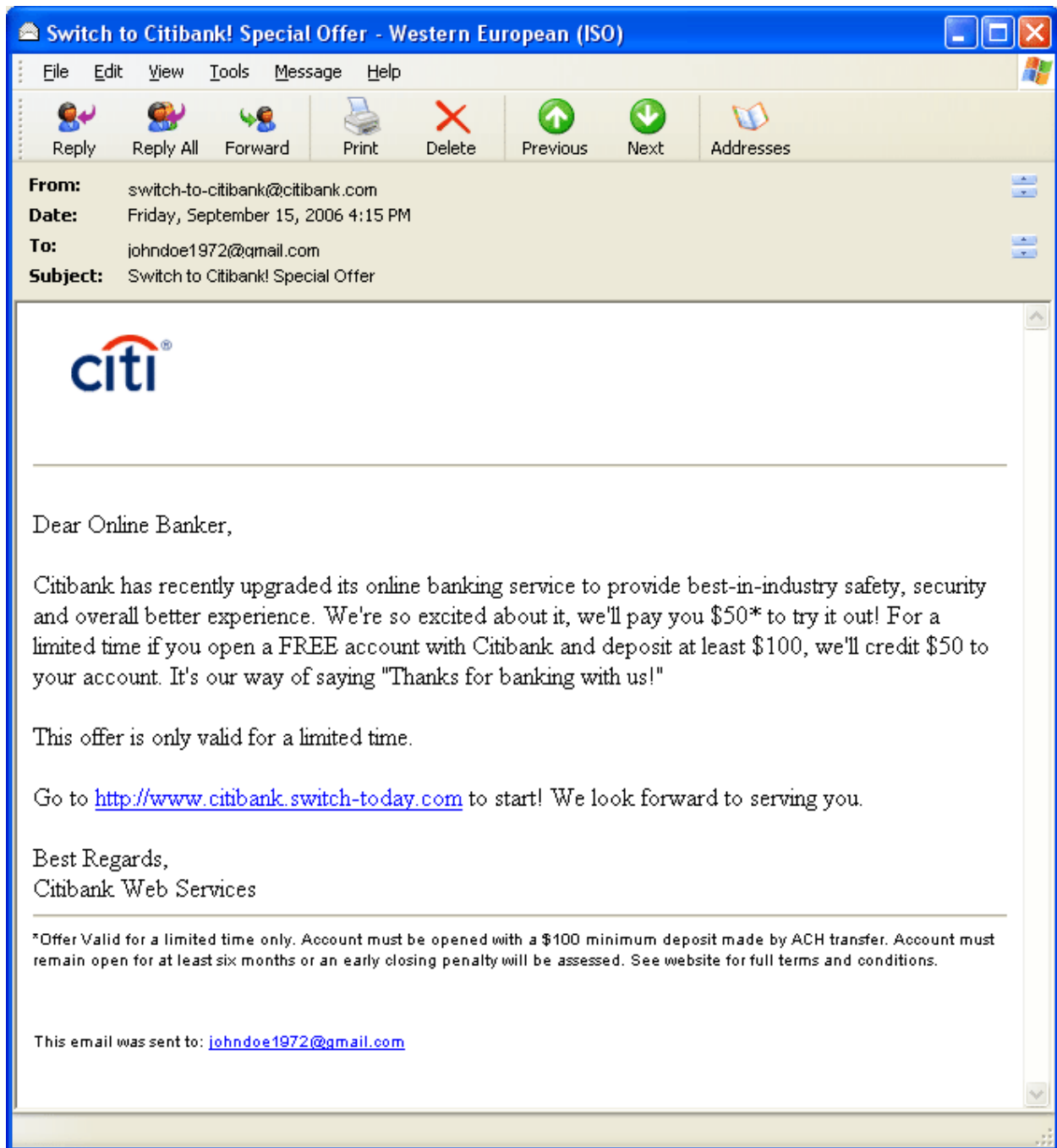


Figure 4: A stimulus from the study by Tsow et al. [44], demonstrating how phishers might devise deceit approaches to take advantage of advertising campaigns, and to target large group of users. Note that this email is relevant to all users who *are not* clients of the impersonated brand, as opposed to those who *are*. If a recipient follows a link, he may be asked to give information relating to his current bank, in order to initiate a transfer of funds to the account he is led to believe he is opening. The stimulus is synthetic, i.e., has not occurred in the wild.

### 3 Server-Side Security

Whereas client-side tools like spam filters and anti-virus software have obvious benefits, we have argued extensively that there are also good reasons not to rely entirely on clients and their machines for their security. Users may fail to see warnings or signs of fraud, and may be duped into installing software that circumvents their malware protection or which plainly performs keylogging, screen scraping or other actions associated with spyware. It has been estimated [32] that 11% of all networked computers run botnet software, and a larger portion still are affected by various forms of malware. Therefore, it is important to complement client-side countermeasures with server-side techniques. We will review some such techniques herein.

#### Remote-Harm Detection

Most phishing attacks today are based either on luring users to enter credentials on sites mimicking legitimate sites, or on installing spyware (such as keyloggers) on their machines. There are many efforts to prevent such threats. Remote-Harm Detection (or RHD) [24] is a recently proposed technique to detect when such an attack has already been performed. Its aim, in other words, is to detect client exposure to phishing and malware attacks *after the fact* and to mitigate the harm that these attacks cause Internet users. The use of reactive defenses is not intended to replace proactive measures, but to provide a second line of defense.

Broadly speaking, RHD involves a web server learning whether a client browser has initiated dangerous Internet connections. Such connections can include browsing of known phishing sites as well as browsing that is indicative of malware infection. Note that for some malware, evidence of infection is best obtained *indirectly*, e.g., by detecting whether the client has been directed to certain ad-bearing web sites or has been attacked by *other* malware relying on similar vulnerabilities. This is an important observation, as it suggests that one can potentially detect the presence of sophisticated malware that takes direct action to erase its tracks.

Depending on the information that an RHD-enabled server gleans about suspected malware and on the transactional relationship between the server and a client, the server or its administrators can initiate defensive action on behalf of the client/user. For example, an Internet bank might reject transactions conducted by customers through clients with probable malware infections, and might take additional measures like contacting vulnerable customers by mail or telephone.

It is worth pointing out that RHD requires no user involvement and no client-side software. In other words, the method relies exclusively on server-side implementation. The advantages of this approach are that the system is transparent to users and avoids the complications and burden of user-mediated software installation. Furthermore, it does not stoke the dangerous habit in users of downloading potentially troublesome executables or patches from the Internet.

The basis for RHD system is a web-browser feature that is referred to as *URL probing*. Most web browsers are configured to retain what is called a *history*, a list of the URLs visited by a user in the recent past—typically the last nine days. Thus, a server can ascertain the presence of a particular URL in the browser of a visiting client. For example, a server can determine if a visiting client has browsed the site `www.xyz.com` in the past few days. It is important to note that URL probing only reveals information in the case of *exact matching*. For example, if a client has visited `www.xyz.com/nihil`, a server will not learn this fact by probing the client's browser history for the URL “`www.xyz.com`.”

**Example: `downloader.trojan`.** As an example of how an RHD system might work, let us consider a Trojan program that is subject to drive-by installation, namely the executable referenced by the antivirus community as `downloader.trojan`. This Trojan, released in 2002, is relatively old. Computers that lack up-to-date virus protection, however, are still vulnerable. This Trojan used to have the ability to install itself from a web site without the user's permission,

making use of a security flaw in the Internet Explorer browser. Recent browser versions require user consent for the program to install. However, as Edelman has pointed out [10], there are deceptive ways of making users install software. Having been installed, this piece of malware downloads and installs other programs from other web sites without the user’s knowledge, and adds bookmarks (also referred to as *favorites*) to the user’s browser.

There are two strategies for an RHD system to perform URL probing to detect the presence of `downloader.trojan` on a client. The sensor can attempt to detect browsing behavior that may have resulted in infection by the Trojan, or else it can search for browsing behavior characteristic of existing infection. In particular, the sensor can probe for:

1. **Browser visits to malware-installing sites:** The server checks to see if the user has visited any sites that are known to attempt to install the Trojan. Note that this may work only for a short time subsequent to infection, namely the length of time that the user’s browser retains entries (usually nine days).
2. **Browser visits that imply malware infection:** The server checks to see if the user has visited sites that match the Trojan’s *signature*—that is, sites that the Trojan, once installed, causes a client to visit and/or bookmark. The presence of a cluster of such sites is indicative of infection.

A naïve deployment of either approach would involve the RHD server harvesting highly specific information about user browsing habits. However, this raises considerable privacy concerns. While acknowledged traces of malware might not be privacy-sensitive for users, some of the URLs associated with malware are also sites that receive actual user-originated traffic (as opposed to traffic resulting from malware infection). This is particularly so in the case of `downloader.trojan`, which bookmarks large numbers of sites with pornographic content. However, as described in [29], it is possible to deploy an RHD system that filters data on the client machine and returns only a binary result. For example, we can design an RHD system that returns only a rough “positive” or “negative” classification of the perceived probability of infection of a given client by `downloader.trojan`, and no information about which sites reside in the client’s browser history. Therefore, if a user has not been infected by a given piece of malware, but has visited sites that happen to be part of the footprint of the malware piece, then a diagnostic should return the response “no infection”, and not “partial match”. On the other hand, if the user’s browser has evidence of visits/bookmarks to *all* the sites corresponding to a footprint, then it is safe to assume that the user’s machine has been infected. (Moreover, even if this were not the case for a small set of users, it would not be stigmatizing for them to be incorrectly classified as having been exposed malware. This is a very unlikely event, though, given the typical sizes of footprints.)

## Takedown vs. Take-home

One way to determine the likely success rates of various phishing attacks, learn the demographics of vulnerable clients, and provide educational feedback to apparent would-be victims of attacks is to change from takedown of offending sites to what we may call “*take-home*”. Take-home works by having all traffic to a phishing site forwarded to an external site that the attacked brand controls, while takehome works by removing access to the phishing site. The external site that receives the redirected traffic would be hosted at the impersonated domain, and therefore, it would receive any cookies that were previously set, since these are automatically<sup>2</sup> released to the site. This helps build an understanding of who is vulnerable to attacks. It also allows the number of accesses to be counted as a function of time; this allows the financial institution to

---

<sup>2</sup>If cache cookies [23] are used instead of conventional cookies, then they would not be automatically sent by the client computer, but would have to be requested from the client computer by the server of the site that traffic is redirected to.

determine what phishing campaigns are the most threatening, and estimate the likely number of actual victims before take-home was initiated. (See [21] for information on the time-dependency of the yield of a given phishing attack for an example of how partial information of number of accesses can help estimate previous accesses and compromises.) In addition, the site may inform the visitor of the risks of phishing, how to avoid it, and of the commitment to security of the financial institution.

A passive version of take-home would simply monitor accesses to graphics, and detect instances where logos and other images are accessed by a party at a given IP address, but where the supporting html documents are not downloaded from the same party. This is an indication of an attack in which the phisher provides a webpage in which credentials are requested, but in which graphics are not stored on the machine controlled by the server but taken directly from the site of the brand. Again, this can be used to estimate the number of victims in ongoing attacks, and to determine the identities and demographics of the clients who visited the site.

## Sticky cookies

A conventional computer cookie is a piece of information stored in a specially designated cache in a Web browser. Cookies can include user-specific identifiers, or personal information about users (e.g., this user is over 18 years of age). Service providers typically employ cookies to personalize Web pages. For example, when Alice visits the Web site X, the domain server for X might place a cookie in Alice's browser that contains the identifier that uniquely identifies Alice. When Alice visits X again, her browser releases this cookie, enabling the server to identify Alice automatically. However, many users (and some software packages) regularly clear their cookies, which frustrates efforts to distinguish legitimate access from attacks.

A cache cookie [23], by contrast, is not an explicit browser feature. It is a form of persistent state in a browser that a server can access in unintended ways. There are many different forms of cache cookies; they are byproducts of the way that browsers maintain various caches and access their contents. Their main benefit is that they are "sticky"—they are harder to remove, and are not cleared along with conventional cookies. Like conventional cookies, they can store identifiers that allows servers to recognize repeat visitors. This allows for better recognition of users, which can be used for purposes of mutual authentication. This is meaningful in the context of approaches such as SiteKey [39]. Another way to authenticate the service provider to the client would be to automatically fill the user name in a form. If consumers become accustomed to having a website autofill their user name, then they may become suspicious of sites that do not. Of course, this is a double-edged sword, as the absence of the user name will be the typical experience of roaming users (who use different computers for different accesses to a legitimate site). Also, it does not address the problem of users who choose their email address as their user name, or any other information that is easily guessed by an attacker.

## Avoiding automated configuration of attacks

As demonstrated on [22], it is possible for attackers to determine the banking affiliation of a given potential victim with a reasonable success probability, without even interacting with the human user. This is achieved by inspection of the browser history of the potential victim, concluding that a person banks with a given institution if he or she has visited their website recently. The attack demonstrated on [22] assumes two rounds of interaction: A first round, in which the affiliation is inferred by a website that the user visits, and a second round in which an appropriately configured email is sent to the user. The attack can also be performed in one round, by combining the two rounds described above [20]. This can be done by an attacker who would send out emails that contain a large selection of logos and other graphics, where these images are stacked in an order that depends on the contents of the browser history. This would be done in a manner so that the graphics of a visited site shows up on top of the other

images (thereby suitably hiding the latter.) This would cause the email to appear to originate exactly from the institution the recipient is banking with. We note that the one-round version of the attack only appears to work against people reading their email using webmail; these people, though, are currently a majority of computer users. Countermeasures to this attack are described in [26, 19], whether on the server side or client side.

## 4 Anticipating Threats from Strengths and Weaknesses

The way phishing attacks will be carried out in a few years is likely to be affected by the strengths and weaknesses of current countermeasures, including law enforcement efforts, technology deployment, and education. We argue that it is possible to further anticipate trends by building hypotheses of likely developments. We give two examples of how we believe this can be done.

### Better detection of spoofed messages

Assume that general spoof detectors<sup>3</sup> will become increasingly successful at blocking attempts at spoofing. We argue that if this happens, then phishers will come to rely more on real domains that they control, and that they increasingly mount attacks based on cousin domains (e.g., `switch-to-citi.com`) and sub-domains (e.g., `bankofamerica.image-update.com` and `jpm.organchase.com`). To proactively defend against attacks in such a scenario, brands can register domains that appear suitable for phishing.

If spoof detection tools become increasingly powerful, we also have to consider the possibility that attackers take another approach that is not based on impersonating brands. One likely approach would be malware-based, e.g., a keylogger. This could be installed either by methods involving deceit (see, e.g., [10, 40]) or using technical vulnerabilities. To counter this threat, brands can make an effort to assure that clients have up-to-date anti-virus software. The latter has to be done in a careful manner, though, to avoid that clients become so focused on anti-malware initiatives that they can be deceived to download “patches and updates” supplied by an attacker impersonating the financial institution. Namely, a client who worries too much about malware and having up-to-date anti-virus software may be deceived to install software that while claiming to be a countermeasure against spyware is, in fact, the spyware itself. That has become increasingly common as of recent.

### Faster ISP takedown

Takedown is one of the most potent countermeasures against phishing attacks today. As the speed of takedown increases, phishers are likely to mount attacks that defeat this measure. One such attack, referred to as a *Distributed Phishing Attack* [27], uses per-victim personalization of Web hosting, and hosts websites on machines controlled by the phisher. In a distributed phishing attack, takedown of an individual server only prevents a small number of victims (or one in the extreme case) from responding with their credentials, as different sets of potential victims are pointed to different sites. The websites can be hosted on botnets, whose recent commoditization has made it practical to deploy thousands of fraudulent Web hosts to collect personally identifying information from their victims. Such machines could be home routers and access points, which—while not able to serve large amounts of traffic (which would not be needed)—do have near-constant connectivity. See [43, 45, 41] for a discussion of such threats.

The use of botnets as hosts for phishing webpages will to a large extent require URLs that are IP addresses. To avoid detection of this, phishers may use spoofing tricks for the address bar

---

<sup>3</sup>Examples of spoof detectors include Domainkeys [6], SenderId [37] and anti-spoof toolbars such as [3, 4, 9, 8, 16, 11]. Also see [5] for an analysis of these toolbars.

[36]. This can be accomplished by using JavaScript to create a new browser window without the address bar (a browser feature), which accesses the phishing site. The phishing site could use JavaScript and CSS to create a frame which hovers at the top of the page (where the real address bar used to be), and which looks like a real address bar. It is even possible for a cleverly-written spoofed address bar to have buttons that work, and to accept keyboard input. Of course, the phisher may display *any* URL in the spoofed address bar, regardless of the URL actually being visited. Another approach would be to use a pharming attack; see [33, 41] for examples of new breeds of such attacks.

## Speaker biography

Dr. Markus Jakobsson is an Associate Professor at Indiana University at Bloomington, Associate Director of the Center of Applied Cybersecurity Research, and a founder of RavenWhite Inc. He is the inventor or co-inventor of over fifty patents, has served as the vice president of the International Financial Cryptography Association, and is a research fellow of the Anti-Phishing Working Group. Prior to his current position, he was Principal Research Scientist at RSA Laboratories, Member of Technical Staff at Bell Laboratories, and Adjunct Associate Professor at New York University. He is an editor of the International Journal of Applied Cryptography and a group editor of the ACM Mobile Computing and Communications Review. He is also an editor of Phishing and Countermeasures (Wiley, 2006), and editor/co-author of upcoming books on crimeware (Symantec Press, 2007), click-fraud (Morgan and Claypool, 2007), and cryptographic protocols (Addison-Wesley, 2007). He has served as the editor of the RSA CryptoBytes for several years. Professor Jakobsson researches fraud, social engineering and phishing, and the prevention of these attacks. He has laid the foundations to the discipline of how to perform experiments to assess risks arising from socio-technical vulnerabilities in the context of current and potential future user interfaces. He consults to the financial industry, and heads the efforts at [www.stop-phishing.com](http://www.stop-phishing.com). More details are available at [www.markus-jakobsson.com](http://www.markus-jakobsson.com).

## Acknowledgments

Many thanks to Sid Stamm for helpful feedback on earlier versions of the paper. Thanks to Alex Tsow, Ankur Shah, Eli Blevis, Youn-kyung Lim, Ari Juels and Jacob Ratkiewicz for permitting extensive citations of material of theirs.

## References

- [1] M. Arata, Jr., “Preventing Identity Theft For Dummies,” Wiley, 2001.
- [2] R. Akavipat, V. Anandpara, A. Dingman, C. Liu, D. Liu, K. Pongsanon, H. Roinestad, M. Jakobsson, “Phishing IQ Tests Measure Fear, not Ability,” To appear in USEC ’07.
- [3] N. Chou, R. Ledesma, Y. Teraguchi, D. Boneh and J. C. Mitchell, “Client-side defense against web-based identity theft,” 11th Annual Network and Distributed System Security Symposium (NDSS ’04), San Diego, February, 2004.
- [4] Cloudmark, Inc. Accessed: January 8, 2007. [www.cloudmark.com/desktop/download/](http://www.cloudmark.com/desktop/download/)
- [5] L. Cranor, S. Egelman, J. Hong, and Y. Zhang, “Phinding Phish: An Evaluation of Anti-Phishing Toolbars,” November 13, 2006. CMU-CyLab-06-018.
- [6] Domainkeys. Accessed: January 8, 2007. [antispam.yahoo.com/domainkeys](http://antispam.yahoo.com/domainkeys)
- [7] R. Dhamija, J.D. Tygar, M. Hearst, “Why Phishing Works,” In the Proceedings of the Conference on Human Factors in Computing Systems (CHI2006), 2006

- [8] EarthLink, Inc. EarthLink Toolbar. Accessed: January 8, 2007. [www.earthlink.net/software/free/toolbar/](http://www.earthlink.net/software/free/toolbar/)
- [9] eBay, Inc. Using eBay Toolbars Account Guard. Accessed: January 8, 2007. [pages.eBay.com/help/confidence/account-guard.html](http://pages.eBay.com/help/confidence/account-guard.html)
- [10] B. Edelman. “How VeriSign Could Stop Drive-By Downloads.” Accessed: January 12, 2007. [www.benedelman.org/news/020305-1.html](http://www.benedelman.org/news/020305-1.html)
- [11] Google, Inc. Google Safe Browsing for Firefox. Accessed: January 8, 2007. [www.google.com/tools/firefox/safebrowsing](http://www.google.com/tools/firefox/safebrowsing)
- [12] A .Emigh, “Mis-Education,” In *Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft*, M. Jakobsson and S. A. Myers (editors). ISBN 0-471-78245-9, Hardcover, 739 pages, December 2006, pp. 260–273.
- [13] P. Finn and M. Jakobsson, “Designing and Conducting Phishing Experiments,” To appear in IEEE Technology and Society Magazine, Special Issue on Usability and Security, 2007.
- [14] B.J. Fogg, C. Soohoo, D.R. Danielson, L. Marable, J. Stanford, E.R. Tauber, “How Do Users Evaluate the Credibility of Web Sites?: A Study with Over 2,500 Participants,” Proceedings of DUX(2003)
- [15] B.J. Fogg, J. Marshall, O. Laraki, A. Osipovich, C. Varma, N. Fang, J. Paul, A. Rangnekar, J. Shon, P. Swani, M. Treinen, “What Makes Web Sites Credible?: A Report on a Large Quantitative Study,” Proceedings of CHI (2001), pp. 61–68
- [16] GeoTrust, Inc. TrustWatch Toolbar. Accessed: January 8, 2007. [toolbar.trustwatch.com/tour/v3ie/toolbar-v3ie-tour-overview.html](http://toolbar.trustwatch.com/tour/v3ie/toolbar-v3ie-tour-overview.html)
- [17] V. Griffith and M. Jakobsson. “Messin’ with Texas, Deriving Mother’s Maiden Names Using Public Records.” ACNS ’05, 2005.
- [18] IP2Location. Accessed: January 8, 2007. [www.ip2location.com](http://www.ip2location.com)
- [19] C. Jackson, A. Bortz, D. Boneh, J. C. Mitchell, “Web Privacy Attacks on a Unified Same-Origin Browser,” In the 15th Annual World Wide Web Conference (WWW06), 2006.
- [20] C. Jackson, personal communication.
- [21] T. Jagatic, N. Johnson, M. Jakobsson, F. Menczer. “Social Phishing.” To appear in the Communications of the ACM, 2007
- [22] M. Jakobsson, T. Jagatic, and S. Stamm, “Phishing for Clues: Inferring Context Using Cascading Style Sheets and Browser History” Accessed: January 8, 2007. [browser-recon.info](http://browser-recon.info)
- [23] M. Jakobsson, A. Juels, T. Jagatic, “Cache Cookies for Browser Authentication (Extended Abstract),” IEEE S&P ’06
- [24] M. Jakobsson, A. Juels, J. Ratkiewicz, “Remote-Harm Detection.” Accessed: January 8, 2007. [rhd.ravenwhitedevelopment.com](http://rhd.ravenwhitedevelopment.com)
- [25] M. Jakobsson and J. Ratkiewicz, “Designing Ethical Phishing Experiments: A study of (ROT13) rOnl auction query features.”, WWW ’06
- [26] M. Jakobsson and S. Stamm. “Invasive Browser Sniffing and Countermeasures.” WWW ’06
- [27] M, Jakobsson and A. Tsow, “Making Takedown Difficult,” In *Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft*, M. Jakobsson and S. A. Myers (editors). ISBN 0-471-78245-9, Hardcover, 739 pages, December 2006, pp. 461–467.
- [28] M. Jakobsson, A. Tsow, A. Shah, E. Blevis, Y.-K. Lim, “What Instills Trust? A Qualitative Study of Phishing,” To appear in USEC ’07.

- [29] A. Juels, M. Jakobsson, J. Ratkiewicz, “Remote-Harm Detection,” manuscript in preparation.
- [30] P. Kumaraguru, Y.W. Rhee, A. Acquisti, L. Cranor, J. Hong, E. Nunge, “Protecting People from Phishing: The Design and Evaluation of an Embedded Training Email System,” Carnegie Mellon University, CyLab, CMU-CyLab-06-017 (November 2006)
- [31] SonicWALL Phishing IQ Test II. Accessed: January 8, 2007.  
[www.sonicwall.com/phishing/](http://www.sonicwall.com/phishing/)
- [32] J. Markoff, “Attack of the Zombie Computers Is Growing Threat,” New York Times, January 7, 2007.
- [33] M. Meiss, “Race Pharming,” In *Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft*, M. Jakobsson and S. A. Myers (editors). ISBN 0-471-78245-9, Hardcover, 739 pages, December 2006, pp. 133–136.
- [34] S. Orr, M. Jakobsson, “Spotting Copycats,” manuscript in preparation.
- [35] Out-law.com, “Staff reveal passwords for a chocolate bar,” [www.out-law.com/page-4469](http://www.out-law.com/page-4469)
- [36] A. Raskin, “Simulated Browser Attack,” In *Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft*, M. Jakobsson and S. A. Myers (editors). ISBN 0-471-78245-9, Hardcover, 739 pages, December 2006, pp. 89–101.
- [37] SenderId. Accessed January 8, 2007.  
[www.microsoft.com/mscorp/safety/technologies/senderid/default.aspx](http://www.microsoft.com/mscorp/safety/technologies/senderid/default.aspx)
- [38] E. Shanahan, “ID Thieves’ New Tricks,” Reader’s Digest (June 2006) pp. 82–87
- [39] SiteKey at Bank of America. Accessed January 8, 2007.  
[www.bankofamerica.com/privacy/sitekey](http://www.bankofamerica.com/privacy/sitekey)
- [40] S. Stamm, M. Jakobsson, “Social Malware,” Experimental results available at [www.indiana.edu/~phishing/verybigad/](http://www.indiana.edu/~phishing/verybigad/)
- [41] S. Stamm, Z. Ramzan and M. Jakobsson, “Drive-By Pharming,” Indiana University Technical Report TR641, December, 2006.  
[www.cs.indiana.edu/pub/techreports/TR641.pdf](http://www.cs.indiana.edu/pub/techreports/TR641.pdf)
- [42] S. Srikwan and M. Jakobsson, “Using comics to explain security threats,” manuscript in preparation.
- [43] A. Tsow, “Phishing with Consumer Electronics – Malicious Home Routers,” In Models of Trust for the Web, a workshop at the 15th International World Wide Web Conference (WWW2006), May 22-26, Edinburgh, Scotland.
- [44] A. Tsow, M. Jakobsson, E. Blevis, Y.-K. Lim, “Deceit and Design: A Large User Study of Phishing”, manuscript in preparation.
- [45] A. Tsow, M. Jakobsson, L. Yang and S. Wetzel, “Warkitting: the Drive-by Subversion of Wireless Home Routers,” Anti-Phishing and Online Fraud, Part II Journal of Digital Forensic Practice, Volume 1, Special Issue 3, November 2006
- [46] T. Whalen and K.M. Inkpen, “Gathering evidence: use of visual security cues in web browsers,” In Proceedings of the 2005 Conference on Graphics Interfaces, pp. 137–144.
- [47] A. Whitten, J.D. Tygar, “Why Johnny Can’t Encrypt: A USability Evaluation of PGP 5.0,” 8th Usenix Security Symposium, 1999, pp. 169–184.
- [48] M. Wu, R. Miller, S. Garfinkel, “Do Security Toolbars Actually Prevent Phishing Attacks?,” In the Proceedings of CHI, 2006