Reinforcing bad behaviour: the misuse of security indicators on popular websites

Douglas Stebila
Information Security Institute, Queensland University of Technology
Brisbane, Queensland, Australia
stebila@qut.edu.au

ABSTRACT

Before making a security or privacy decision, Internet users should evaluate several security indicators in their browser, such as the use of HTTPS (indicated via the lock icon), the domain name of the site, and information from extended validation certificates. However, studies have shown that human subjects infrequently employ these indicators, relying on other indicators that can be spoofed and convey no cryptographic assurances. We identify four simple security indicators that accurately represent security properties of the connection and then examine 125 popular websites to determine if the sites' designs result in correctly displayed security indicators during login. In the vast majority of cases, at least some security indicators are absent or suboptimal. This suggests users are becoming habituated to ignoring recommended security indicators.

KEYWORDS

user education, security indicators, web browsers, HTTPS

CATEGORIES AND SUBJECT DESCRIPTORS

H.5 [**Information Interfaces and Presentation**]: Miscellaneous; K.4.4 [**Computers and Society**]: Electronic Commerce—*security*

INTRODUCTION

Users on the Internet are regularly confronted with complex security decisions that can affect their privacy. They must decide whether it is safe to enter their username, password, credit card details, and other personal information on websites with very different interfaces and only a few visual clues on whether it is safe to do so. These *security indicators* include the protocol used, the domain name, the SSL certificate, and visual elements in the browser window. Very few users understand the technical details of these various indicators. Not surprisingly, users often get it wrong, either ignoring security checks completely or misunderstanding them.

We begin by identifying four security indicators, present in all major web browsers, that together convey strong security assurances. We then examine 125 top websites, including major banks in several countries, to ascertain the security indicators present at log in. We find that the login interaction sequence for the vast majority of sites results in missing or incorrect security indicators. While several studies have evaluated whether users correctly use security indicators, there has been no work investigating why that is so. Our data supports the hypothesis that users do poorly at understanding security indicators because they are effectively trained to do so by website designers. Ours is the first study assessing a variety of security indicators on popular websites from this perspective.

OZCHI 2010, November 22-26, 2010, Brisbane, Australia

OZCHI 2010 Proceedings ISBN: x-xxxxx-xxx-x

SSL/TLS AND SECURITY INDICATORS

The Secure Sockets Layer (SSL) protocol was proposed in 1995 by Netscape for securing web and Internet traffic; later versions of SSL were renamed the Transport Layer Security (TLS) protocol, and here we refer to them interchangeably. The use of TLS to secure web content delivered by the HyperText Transport Protocol (HTTP) is commonly referred to as the HTTPS protocol.

TLS provides confidentiality, message integrity, and entity authentication. Typically, authentication is based on public key certificates which are issued by a *certificate authority* (CA) and which bind an identifier – typically a domain name – to a public key. Before issuing a certificate, a CA is often presumed to perform some offline validation checks to confirm the entity requesting the certificate is legitimately related to the domain name in question. Recently, CAs have started issuing *extended validation* (EV) certificates which, in theory, have a more thorough validation process (and, in practice, cost more). Nonetheless, there is debate over the effectiveness of EV certificates and whether they achieve their goals.[1]

When a user's web browser establishes a TLS connection with a website, the browser performs a number of checks on the website's certificate (for example, it matches the website's domain name, is signed by a CA the user's web browser trusts, and has not expired). Once the checks pass and a connection is successfully established, web browsers typically display additional information when communicating with sites using TLS. These are referred to as *security indicators*.[2] Internet security recommendations from companies and governments typically advise users to check for these security indicators before entering sensitive personal information.

We reviewed the academic literature and security recommendations from online businesses to identify potential security indicators. We identified the following four security indicators that are actually indicative of the strong security offered by HTTPS/TLS.

[1]For example, one of the sites in our study had an EV certificate issued to a company with the very descriptive name "Various Inc."
[2]We exclude TLS-related warnings, such as expired certificates or unknown CAs, from the definition of security indicators. We only include indicators present *after* the TLS connection is established.

SI1. Does the URL in the location bar begin with https? [12, 13, 9, 3, 4]
SI2. Is the domain name of the URL in the location bar correct? [16, 14, 15]
SI3. Is the lock icon displayed somewhere in the browser chrome (the non-content part of the browser window, such as the toolbar, status bar, title bar, or location bar)? [16, 12, 13, 9, 3, 15, 4]
SI4. Are there indicators present for an extended validation certificate, such as a green background in the location bar and the name of the company on the certificate? [10]

However, many studies have shown that users do not properly check these recommended security indicators.

Firstly, users often ignore the recommended security indicators. Whalen and Inkpen [16] used eye tracking data to find that users frequently look at the lock icon during log in but stop looking for security information after logging in. Schechter et al. [12] removed security indicators from financial websites but all 63 users in their study continued to enter their password. Sobey et al. [13] found (using eye tracking data) that indicators associated with EV certificates did not change user behaviour very much, even though they are purported to offer better authentication promises. Sunshine et al. [14] observed that around 30% of users ignore TLS warnings for expired certificates, unknown CAs, and domain name mismatches, and a majority do not understand these warnings.

Secondly, users often incorrectly use non-secure properties of web pages to make security decisions. Dhamija and Tygar [6] found that users cannot reliably distinguish browser chrome from web page contents, so they often consider images displayed on a web page to be trustworthy. This is troubling considering the threat of *visual spoofing* [1] in which a web page disables certain parts of the browser chrome (such as the location bar) and replaces those regions with spoofed HTML versions that display misleading information (such as a fake location bar with a falsified domain name). Users also have been reported to make decisions based on the type of site [9, 16], the type of information being submitted [9, 16], and text on the page that says the site is secure [16]. It is of significant concern that some design literature [7, §5.2.1] *recommends* the use of design elements that "maximise perceived trustworthiness" but in fact further conflate where to look for security indicators.

Some websites have deployed *site authentication images*, which are custom images, selected by users at registration time, that are displayed to users with the recommendation that users stop if they do not see the correct site authentication images. Unfortunately, site authentication images have no real security: they offer no cryptographic security properties and are vulnerable to man-in-the-middle attacks. Schechter et al. [12] found that 92% of users proceeded even when their site authentication image was absent. Most disturbingly, some of the users in that study admitted to ignoring TLS-related warnings *because* site authentication images were present.

We note a contrast between our work, which evaluates the presence of security indicators on popular websites, and a recent survey [4], which evaluates password login practices on popular websites.

STUDY OVERVIEW

We constructed a list of popular websites consisting of:

- the top 110 domain names by traffic, as ranked by Alexa Topsites[3] [2], excluding country-specific duplicates (e.g., after using google.com, exclude google.co.uk) but including distinct properties or brands (e.g., still include youtube.com),
- the top 10 banks in the USA [8], the "big 4" banks in the UK and Australia, and the "big 5" banks in Canada, and
- the top 4 North American webmail providers (Yahoo, Microsoft, Google, and AOL).

Over a two-week period in September 2009, we visited each site to determine the security indicators present when users attempt to login to the site. Our experimental procedure was as follows:

1. Clear all user data (cache, cookies); restart the web browser (Mozilla Firefox 3.5).[4]
2. Visit the site by typing the domain name as listed in the Alexa Topsites ranking and hitting enter.
3. Find the page containing the login form, first by looking on the front page of the site, and then looking for "Log in" or "Sign in" links. (For 34 pages not in English, we used Google Translate to help us find links to the login form.) Stop if no login is found.
4. Record data about the login page: the URL, presence of the location bar, protocol of the page (HTTP or HTTPS), use of EV certificates, protocol of the login form submission, and presence of lock icons on the page or as the favicon.

Of the 137 websites we visited, 12 provided no method for logging in and are excluded from the rest of this study. We classified the types of the remaining 125 sites in Table 1.

Type of site	Examples	n
bank	bankofamerica.com, anz.com	23
blogging / hosting	blogger.com, wordpress.com	10
content	youtube.com, wikipedia.org	17
corporate	microsoft.com, adobe.com	4
e-commerce	ebay.com, amazon.com	6
file sharing	flickr.com, rapidshare.com	8
mail	mail.yahoo.com, hotmail.com	5
pornography		7
portal / search	google.com, yahoo.com	29
social networking	facebook.com, myspace.com	16

Table 1: Classification of sites visited.

MISUSE OF SECURITY INDICATORS

We identified a number of web page design characteristics that make it hard to correctly evaluate the recommended security indicators. This misuse of security indicators is discussed in detail in the remainder of this section and summarized in Table 2.

Misused security indicator	n
HTTP login page with HTTPS form submission	19
Lock icon on page or as favicon	29
Hidden location bar	2
Mismatched domain name	16
Very complicated URL	53

Table 2: Summary of observations.

[3]One limitation of the Alexa Topsites data is that it ranks the most frequently visited websites, not the most frequently logged-in-to websites. For example, Wikipedia is 6th on the list but few users log in to Wikipedia; on the other hand, Facebook is 2nd on the list and one cannot use Facebook without logging in. While a dataset reflecting the most frequently logged-in-to websites would be very helpful, we do not know of one.

[4]Although different browsers display security indicators slightly differently, they do display effectively the same security indicators, so our choice of browser does not affect the experimental results.

HTTP login page with HTTPS submission

For SI1 and SI3 to be satisfied, the login page and all elements on the page must be delivered using HTTPS. For the user's credentials to be protected, the login form must submit the data using HTTPS. We recorded in Table 3 the protocol used to deliver the login page and compared it with the protocol with which the login form data was submitted.

Login page	Form submission	n
HTTP	HTTP	56
	HTTPS	13
	HTTPS w/EV cert.	6
HTTPS	HTTPS	40
HTTPS w/EV cert.	HTTPS w/EV cert.	10

Table 3: Protocol of login page compared to form submission.

Of course the most troubling piece of data is that 45% of sites submitted usernames and passwords with no encryption whatsoever.

Several sites (19, or 15%) delivered the login page over HTTP but submitted the login form using HTTPS. While it is good that these sites protect user data in transit, users do not know this a priori unless they view the HTML source of the page. Thus, a security-conscious user should not login to these sites, even though the login details will probably be protected. Troublingly, Facebook (the 2nd ranked site on Alexa Topsites) follows this practice.

In some cases this design decision can be understood since the login form appears on the main page and delivering the main page over HTTPS may result in an unacceptably high server load. However, for other sites, such as Facebook, the only purpose of displaying the main page is to display the login form which might as well have been delivered over HTTPS since the typical user's next action (logging in) will result in an HTTPS connection anyway, and HTTPS connections can be reused over subsequent requests.

Curiously, 6 of these 19 sites have an EV certificate; since the main purpose of EV certificates is to provide additional security indicators in the browser, it is surprising that website designers purchased an EV certificate at greater cost but do not deliver login pages utilizing these additional security indicators. We also note that only 13% of the sites used EV certificates in any way.

Two sites allowed users to choose on the login page whether to submit the form via HTTP or HTTPS. One site delivered the main page over HTTP but the login form on the main page was included via an `<iframe>` from an HTTPS URL; this does not result in any security indicators being displayed.

Lock icon on page or as favicon

Users are accustomed to looking for a lock (or key) icon to check if the page is secure. In particular, users should check for the lock icon in the browser chrome (SI3), not in the content delivered by the web page. Unfortunately, studies have shown users have trouble distinguishing content from browser chrome [6]. To help users avoid confusion, good website design would preclude the use of lock icons in web page content. Unfortunately, this is not the case.

Of the sites we surveyed, 15% displayed a lock icon on HTTPS pages and an additional 5% displayed a lock icon even on HTTP pages. Notably, 70% of banks displayed a lock icon on pages. This may be related to studies which show that users have greater trust in sites that display statements about the use of SSL, but exacerbates the problem of users conflating trusted and untrusted content.

Most browsers allow websites to display an icon in the location bar or tab bar; this is referred to as the *favicon* since it is stored when users add the page as a bookmark/favourite. While supplied by the website, it is displayed in areas that are conceptually part of the browser chrome, not web page content. We observed 2 sites (both run by AOL) used a key for the favicon when on HTTPS pages.

Hidden location bar

In order to evaluate SI1 (a URL beginning with https) and SI2 (the domain name in the URL being correct), users must examine the URL in the location bar. In all major browsers the site cannot hide the location bar in the main window. However, the `window.open` Javascript function does allow a web page to open new windows with the location bar hidden. (Although Google Chrome and recent versions of Mozilla Firefox no longer allow this.) With the location bar hidden, the user cannot evaluate SI1 or SI2, and an attacker could employ visual spoofing to construct a fake location bar.

Of the 125 sites we surveyed, only 2 employed a login method involving a popup window with a hidden location bar. Curiously, both of these sites were Australian banks, suggesting the possibility of some hidden factor leading to this design (a common subcontractor designed the websites? industry recommendations?).

Further, the security guidelines of one of these banks [3] provides incorrect recommendations for assessing SI1 and SI2, suggesting a misunderstanding of security principles by the website designer. It recommends checking for the website address as displayed in the *title bar* of the popup window. This recommendation rests on the fact that some browsers display the URL in the title bar if the HTML page has no `<title>` tag specified, but is easily spoofed by an attacker who specifies a `<title>` tag of their choice.

Mismatched domain name

For SI2, users are supposed to check if the domain name of the URL in the location bar is correct. This helps avoid phishing and redirection attacks. However, what does "correct" mean? One measure of correctness is whether the domain name of the login page matches the domain name that the user typed. They may differ if the user follows a link or is redirected to another page.

We visited each website by typing in the domain name as found on the Alexa Topsites list (e.g., google.com). We compared the domain name of the login page to the domain name we typed, and recorded in Table 4 whether it was an *exact match* (we allowed the site to prepend the string www. since the use of this prefix is nearly ubiquitous), a *close match* (login.yahoo.com is a close match for yahoo.com), or *no match*.

Match?	Example typed domain → login domain	n
Exact match	google.com → www.google.com	71
Close match	yahoo.com → login.yahoo.com	38
No match	hotmail.com → login.live.com	16

Table 4: Mismatches of typed domain name vs. login page domain name.

The majority of "no match" cases were sites owned by parent corporations but using separate branding for the site; for example, Flickr is owned by Yahoo! and uses Yahoo!'s single sign-on via login.yahoo.com. Other non-matching sites were banks. For example, the Royal Bank of Scotland (rbs.co.uk) has its login page on www.rbsdigital.com. Even though users are redirected by clicking on a link on the main page, they can still be tricked if the main web page is not delivered via HTTPS (it was not) or if the users do not verify the main page's authenticity before clicking on the link.

Complicated URL

Web design experts have long viewed the URL as a form of user interface [11] and eyetracking research confirms that users look at URLs frequently [5].

Since SI1 and SI2 involve reviewing the URL in the location bar, and in particular the first part of the URL (the protocol and domain name), a complex of the URL can make it harder for for users to make correct security decisions. We recorded the full URL of each login page and developed three classifications of complexity:

1. Simple URL: just the domain name; for example, https://www.chase.com/. $n = 50$, average URL length 21.9.
2. Medium complexity URL: one short path component after the domain name; e.g., https://www.blogger.com/start. $n = 22$, average URL length 32.9.
3. Very complicated URL: lots of characters after the domain name; for example, the login page for youtube.com had the 224-character URL https://www.google.com/accounts/ServiceLogin?uilel=3&service=youtube&passive=true&continue=http%3A%2F%2Fwww.youtube.com%2Fsignin%3Faction_handle_signin%3Dtrue%26nomobiletemp%3D1%26hl%3Den_US%26next%3D%252F&hl=en_US&ltmpl=sso (this was not even the longest login URL we saw!). $n = 53$, average URL length 110.1.

The average length of URLs in our study was 56.9 characters; the average length of domain names we typed was 10.9 characters (22.9 characters including the prefix https://www.).

Good web browser design can help here: Google Chrome, for example, displays the domain name in black but the rest of the URL in grey, thereby emphasizing the domain name.

Site authentication images

We identified only 3 uses of site authentication images in the login process, all of which were on sites run by Yahoo!. They displayed the image on the login page which asked for both the username and password. The image was computer-specific.

Seven banks had a multi-stage login procedure, with a first screen asking for a username and then a second screen asking for a password. Some of these sites may be designed to display a site authentication image if the user has one registered for that username; since we did not have usernames at any of these banks, we could not determine if this was the case.

DISCUSSION AND DESIGN RECOMMENDATIONS

Of the 125 sites we evaluated, only 5 avoided all misleading security indicators identified in the previous section. Hence, a typical Internet user will, much more often than not, be asked to make security decisions against best-practice recommendations on security indicators. User education is often identified as a method for increasing security. This paper shows how users are being educated, through daily web use, to ignore recommended security indicators.

If we want users to learn to engage appropriately with security indicators, we need websites where the interaction sequence ensures proper display of security indicators. We encourage website designers to follow these recommendations:

- Deliver the page containing the login form over HTTPS.
- Don't try to hide the location bar.
- Ensure the domain name of the login page matches the domain name of the site in question.
- Don't use lock icons anywhere in the web page content.
- Try to use simple URLs, especially for the login page.

We note that our third recommendation conflicts with single sign-on systems. Given the increasing prevalence of such systems, we believe an important piece of future work is understanding how users assess the redirections and security of single sign-on systems.

This study aimed to determine which security indicators are displayed by popular websites. It remains an important task to design and evaluate better ways for web browsers to display security indicators, including whether browsers should continue to display untrusted favicons in areas considered part of the trusted chrome.

ACKNOWLEDGEMENTS

The author gratefully acknowledges helpful discussions with L. Jean Camp and Ken Radke.

REFERENCES

[1] A. Abelsbach, S. Gajek, J. Schwenk. Visual spoofing of SSL protected web sites and effective countermeasures. In *Proc. Information Security Practice and Experience (ISPEC) 2005*, Springer, 204–216.

[2] Alexa Topsites, Sept. 8, 2009. http://www.alexa.com/topsites

[3] Australia and New Zealand Banking Group Limited (ANZ). Bank safely online, Aug. 2009. http://www.anz.com/personal/ways-bank/internet-banking/protect-banking/bank-safely-online/

[4] J. Bonneau, S. Preibusch. The password thicket: technical and market failures in human authentication on the web. In *Proc. Economics of Information Security (WEIS 2010)*.

[5] E. Cutrell, Z. Guan. What are you looking for?: an eye-tracking study of information usage in web search. In *Proc. CHI 2007*, ACM, 407–416.

[6] R. Dhamija, J. D. Tygar. The battle against phishing: Dynamic security skins. In *Proc. SOUPS 2005*, ACM, 77–88.

[7] F. N. Egger. Affective design of e-commerce user interfaces: How to maximise perceived trustworthiness. In *Proc. Conf. Affective Human Factors Design 2001*, Asean, 317–324.

[8] FDIC. Top 50 commercial banks and savings institutions by total domestic deposits, Jun. 2008. http://www2.fdic.gov/sod/sodSumReport.asp?barItem=3&sInfoAsOf=2008

[9] B. Friedman, D. Hurley, D. C. Howe, E. W. Felten, H. Nissenbaum. Users' conceptions of web security: a comparative study. In *Proc. CHI 2002*, ACM, 746–747.

[10] C. Herley, P. van Oorschot, A. S. Patrick. Passwords: If we're so smart, why are we still using them? In *Proc. Financial Cryptography 2009*, Springer, 230–237.

[11] J. Nielsen. URL as UI, Mar. 1999. http://www.useit.com/alertbox/990321.html

[12] S. E. Schechter, R. Dhamija, A. Ozment, I. Fischer. The emperor's new security indicators: An evaluation of website authentication and the effect of role playing on usability studies. In *Proc. S&P 2007*, IEEE, 51–65.

[13] J. Sobey, R. Biddle, P. van Oorschot, A. S. Patrick. Exploring user reactions to new browser cues for extended validation certificates. In *Proc. ESORICS 2008*, Springer, 411-427.

[14] J. Sunshine, S. Egelman, H. Almuhimedi, N. Atri, L. F. Cranor. Crying wolf: An empirical study of SSL warning effectiveness. In *Proc. USENIX Security 2009*.

[15] Wachovia Security Plus Customer Center – FAQ, Aug. 2009. https://www.wachovia.com/foundation/v/index.jsp?vgnextoid=cd591341de0aa110VgnVCM1000004b0d1872RCRD#9

[16] T. Whalen, K. M. Inkpen. Gathering evidence: use of visual security cues in web browsers. In *Proc. Graphics Interface 2005*, Canadian Human-Computer Comm. Soc., 137–144.