Phishing - Confirming That a Link Is Legitimate

Due to ongoing IT security education and training over the last few years, our vigilance as a community has increased tremendously. This is, of course, a positive development. One side-effect, however, is that the very same vigilance encouraged by that training can sometimes lead us to mistrust legitimate emails.

Phishing emails try (among other things) to trick us into clicking a link and visiting a malicious website. They represent a very real threat. However, sometimes we receive emails from legitimate sources (sometimes third parties contracted by the university) that urge us to click a **legitimate** link.

This document intends to convey a technique for determining the safety of a link included in an email without going into too much detail about the underlying technical principles (DNS hierarchy, website directory organization, etc.).

The Basics - Look For the Usual "Red Flags"

First, apply your own common sense, intuition, and **previously received IT security training** in evaluating the trustworthiness of the received email. If, after doing so, you still feel there's a strong chance the email is legitimate, there is one additional technique that can be used to analyze and verify the link you're being asked to click before actually doing so...

Only After Careful Examination - For Emails You Think Are Likely Legitimate

As an example, we'll be using an actual, legitimate email sent to me asking me to take part in the "UC Cyber Security Awareness Fundamentals" training (screenshot follows further below). Training emails such as these often illicit worry and concern because the return address on them (noreply@sumtotalsystems.com) is not from a UCSB organization or contractor that is readily recognized. So, the first impression given by this email may be that it is not legitimate. And, it should be pointed out that even emails with valid/legitimate return addresses can be malicious (the return address can be spoofed/faked and even non-functional. . .the malicious intent is not that you reply, but that you click a malicious link or open a malicious attachment).

So, in this example, we've been given reason for pause/concern. Yet we suspect it might still be a legitimate business communication. In this situation, it makes sense to judge the intent of the email, which is to get us to click a link, and then analyze/verify the legitimacy of that link. To do so in this and any other similar case, you can:

1. Hover your mouse cursor over (but do *not* click) the link in question. If on a mobile device tap-and-hold (don't just tap) the link in question.

2. Note that at the bottom of your screen/window, the actual destination encoded in the link will appear. (the way this is displayed on mobile devices varies, but it will likely pop up a question asking if you really want to go to the specified site . . . allowing you to opt out of doing so).
Evaluating a Link/Address in Detail

First. . . does the link in the body of the email match the actual destination shown at the bottom of the screen? Malicious actors will often encode a different actual destination than the one presented in the email body.

Second. . . locate the the portion of the address that is between the beginning "http://" part and the very first single forward slash (/). This is the actual address of the server you'll be visiting. The rest (all the verbiage and slashes thereafter) is just subfolders and pages on that server. For example, in the following links, we're interested in only the part in red.

- http://some/otherstuff/mypage.html
- https://eci.engineering.ucsb.edu/accounts/andother/stuff/
- https://www.learningcenter.ucsb.edu/learning/stuff/now/

Note: We're using "http://" and "https://" inter-changeably throughout this document since you're likely to still encounter both. But for our purposes here, it doesn't really matter.

Third. . . to interpret that portion between the http:// and the very first forward slash, it's important to understand that the domain structure is read from right-to-left ("backwards" from the perspective of reading English). So we read back from the first single forward slash towards the http:// portion. We do this to determine what server we'll actually be visiting if we click the link. If we see any names that we don't recognize or trust --reading right-to-left-- before we see the names that we do trust/recognize, it's almost certainly a malicious site. Here is an example...

- https://ucsb.edu.unsafeSite.com/ignore/anything/from/here/because/it/is/after/the/first/single/forward/slash
Reading from right to left (from the first single forward slash), we see "com"... and then "unsafesite"... uh-oh! Of course, we then see "edu" and "ucsb" further to the left. But, those don't matter. The "unsafesite" after the "com" means "unsafesite.com" is where we're going. And we don't want to go there. Here, the owners of "unsafesite.com" have created subdomains called "edu" and "ucsb" so that an unwary potential victim will see "ucsb" and "edu" in the address and think "How can that be malicious!?! It must be safe!" But, the nature of the DNS hierarchy means that you're actually going to visit unsafesite.com.

Here's another example of a malicious address:

- https://www.unsafesite.com/ucsb/edu/tustmeiamsafe.html

Here, a malicious actor has not bothered to create any subdomains, hoping that just having "ucsb" and "edu" farther along in the address (after the first forward slash) will trick the average potential victim. But, because we now know to only look between the "https://" and the first solo "/", and then focus primarily on the portions closest to that first "/", we won't be fooled by that address and it will immediately trigger our "this is unsafe/fake" alarm.

So... what makes "http://www.learningcenter.ucsb.edu/" safe according to the "rules" above? After all, it has "learningcenter" in it? And that could be anybody! The important factor here is that "learningcenter" appears between the https:// and the first / but is also to the left of ucsb.edu. Put another way, reading right-to-left from the first solo forward slash, "ucsb.edu" comes before "learningcenter." Thus, that means that "learningcenter" is under the control of "ucsb.edu." Thus, it is safe.

**When in Doubt, We're Here to Help!**

Though the above can be useful in determining the validity and legitimacy of a link/address found in an email (or elsewhere), we realize that to those not actually interested in the day-to-day functioning of IT, DNS, and, well, the internet... it may not always be that easily remembered. And, with cybersecurity always such a pressing concern, it is normal and natural to feel anxiety about clicking any link even when you're pretty sure it's safe. As always, even with all the above said, if you have any doubt about the legitimacy of an email, please always feel free to ask us about it at help@engineering.ucsb.edu.