We analyzed the entire web and found a cybersecurity threat lurking in plain sight

June 29 2024, by Kevin Saric

Credit: Pixabay/CC0 Public Domain

Our latest research has found that clickable links on websites can often be redirected to malicious destinations. We call these "hijackable hyperlinks" and have found them by the millions across the whole of the

web, including on trusted websites.

Our paper

, published at the 2024 Web Conference, shows that cybersecurity threats on the web can be exploited at a drastically greater scale than previously thought.

Concerningly, we found these hijackable hyperlinks on the websites of large companies, religious organizations, financial firms and even governments. The hyperlinks on these websites can be hijacked without triggering any alarms. Only vigilant—some might say paranoid—users would avoid falling into these traps.

If we were able to find these vulnerabilities across the web, so can others. Here's what you need to know.

What are hijackable hyperlinks?

If you make a typo when entering your bank's web address, you might accidentally end up on a phishing site—one that impersonates, or "spoofs," your bank's website to steal your personal info.

If you're in a rush and don't inspect the website closely, you may enter sensitive personal details and pay a steep price for your mistake. This could include

identity theft, account compromise or financial loss.

Something even more dangerous happens when programmers mistype web addresses in their code. There's a chance their typo will direct users to an internet domain that has never been purchased. We call these phantom domains.

For example, a programmer making a link to theconversation.com might accidentally link to tehconversation.com—note the misspelling. If the mistyped domain has never been purchased, someone could come along and buy that phantom domain for around A$10, hijacking the inbound traffic. In these cases, the price of programmers' mistakes is paid by the users.

These programmer linking errors don't just risk directing users to phishing or spoofing sites. Hijacked traffic can be directed towards a range of traps, including malicious scripts, misinformation, offensive content, viruses and any other hacks the future will bring.

Over half a million phantom domains

Using high-performance computing clusters, we processed the whole browsable web for these vulnerabilities. At a scale never seen in research, in total we analyzed over 10,000 hard drives' worth of data.

Doing so, we found over 572,000 phantom domains. The hijackable hyperlinks directing users to them were found on many trusted websites. In a twist of irony, this even included web-based software designed to enforce privacy legislation on websites.

We investigated what errors caused these vulnerabilities and categorized them. Most were caused by typos in hyperlinks, but we also found another type of programmer-generated vulnerability: placeholder domains.

When programmers develop a website that does not yet have a specific domain, they often enter links to a phantom domain with the expectation the links will be fixed later.

We found this to be common with website design templates, where the

aesthetic components of a website are purchased from another programmer rather than developed in-house. When the design template is later installed on a website, the phantom domains are often not updated, making links to them hijackable.

To determine if hijackable hyperlinks could be exploited in practice, we purchased 51 of the phantom domains they point to and passively observed the inbound traffic. From this, we detected substantial traffic coming from the hijacked links. Compared to similar new domains that lacked hijacked links, 88% of our phantom domains got more traffic, with up to ten times more visitors.

What can be done?

For average web users, awareness is key. Links cannot be trusted. Be vigilant.

For those in charge of companies and their websites, we suggest several technical countermeasures. The simplest solution is for website operators to "crawl" their websites for broken links. Countless free tools are available for doing so. If any broken links are found, fix them before they are hijacked.

We, the Web

British scientist Sir Tim Berners-Lee first proposed the web at CERN in 1989. In his earliest description of it—still widely available on the web as a testament to itself—there is a section titled "non requirements," where security is addressed. This section includes the fateful phrase:

"[Data security is] of secondary importance at CERN, where information exchange is still more important."

While this was true of CERN in 1989, the web is now the primary information exchange medium of the modern age.

We have come to treat the web as an external component of our own brains. This is evidenced by the popularity of large language models like ChatGPT, which themselves are trained on data from the web.

As our dependence deepens, it might be time to mentally re-categorize web data security from "non requirements" to "important requirements."

Provided by The Conversation

Citation: We analyzed the entire web and found a cybersecurity threat lurking in plain sight (2024, June 29) retrieved 9 September 2024 from https://techxplore.com/news/2024-06-entire-web-cybersecurity-threat-lurking.html