

Strider URL Tracer with Typo Patrol

Tony Bradley, CISSP-ISSAP 2006-06-27

Introduction

This article looks at Microsoft's free Strider URL Tracer with Typo-Patrol to help fight typo-squatters and domain parking abuse. The tool can be used to protect children from seeing inappropriate or explicit sites that they should not see, and for companies or trademark owners to scan and investigate sites that may be typo-squatting their domain(s) so that they can be investigated and/or prosecuted.

Rise of the Typo-Squatters

It is inevitable that people surfing the Internet will have a typographical error (typo) every now and again. As one's fingers blaze across his keyboard tapping away, it's easy to hit a key you didn't intend to, or one's left hand will get ahead of his right and might transpose two characters in a word.

Many applications, such as word processors, will instantly recognize the error and underline the misspelled word or suggest an alternative word. Web browsers are not that intelligent or forgiving though. If you are trying to visit `http://www.ebay.com` and you accidentally swap the 'A' and the 'Y' and type `http://www.ebya.com` instead, your Web browser will not alert you to the flaw. It will dutifully take you to the web site registered under the ebya.com domain name.

One might expect that ebya.com would simply result in a standard "Page Not Found" error. However, those lacking in Web ethics have figured out that they could profit from people's typing mishaps. By registering the domain names associated with common typing errors from popular domain names, Web sites can benefit from a steady flow of misguided traffic. The example above, ebya.com, ends up redirected to a Web site at `http://www.megago.com`, which appears to be some sort of index listing of categories of other Web sites to visit.

Many of these typo domains direct users to sites that are inundated with pop-up advertising or possibly even malware such as viruses or worms. Some typo domains of popular children's web sites even redirect to pornographic adult sites, exposing children to inappropriate material because of an innocent mistake typing. Whether it is just a nuisance, a source of malware, or results in exposing minors to adult material, typo domains are a real problem on the Internet.

The business Of domain parking

One might ask himself, "why would someone bother to redirect people to a Web site they don't want to visit?" Logically, you would think that if a person was trying to visit `disney.com` and they enter the name wrong and end up at a different domain, they aren't going to be interested in the product or information on that site. They would simply retype the

domain correctly to go about planning their vacation to Disney World (or continue figuring out how large a second mortgage they need to finance such a vacation).

The reality, though, is that enough people seem to be interested in information on these typo sites. Just as the majority of people abhor email spam and wouldn't think of responding to it or purchasing any products or services promoted with it, enough people will do so that it is still quite profitable.

In most cases, the typo domain is not even selling a product or service itself. The typo domain makes its money from syndicated advertising such as Google's AdSense program. The typo-squatter simply parks the domain and the only content on the site ends up being the ads served from a syndicated advertising program.

With ad syndication, context-sensitive ads are displayed that are based on the overall content of the target web site. When a URL is typed into the address bar or clicked on, the Web browser is instructed to retrieve data from a third-party URL. The third-party URL, using information it knows about the target URL, and possibly combined with details about the user, then serves contextual ads that are relevant to the site or user.

In theory, there is nothing wrong with this practice. If I am visiting a site about golfing, it makes sense that I would want to see advertising that has to do with golfing as well, as opposed to ads about the latest cholesterol drug or mail-order DVD service.

Some domain owners abuse the ad syndication system, however, by simply parking the domains so that the only content on the site to begin with is from the syndicated ads. These sites provide no real value and serve no better purpose than to generate ad revenue for the domain owner. With domain registrations as low as \$7, the domain could pay for itself with as little as one unique visitor every 2 days.

Microsoft's Strider Typo-Patrol

To try to identify and combat this type of systematic typo-squatting and abuse of the syndicated advertising system, Microsoft's Cybersecurity and Systems Management research group developed the [Strider Typo-Patrol tool](#). At the time of this writing, Strider Typo-Patrol works only with Windows XP and Internet Explorer 6. It also requires version 2.0 or higher of [Microsoft's .NET framework](#) before it can be installed. The .NET framework is a 22.4Mb download that is not likely to be installed by default on most home systems, but fortunately it is easy to install.

Components of Strider Typo-Patrol

The Strider Typo-Patrol tool is made up of three major components: typo-neighborhood generator, typo-neighborhood scanner and typo-domain database. The three functions of the Strider Typo-Patrol tool allow users to identify and scan for typo-squatting domains and to contribute to the running list of typo-domains stored in the typo-domain database on Microsoft's servers.

The Strider typo-neighborhood generator takes a given domain, input by the user, and extrapolates all of the conceivable domains that could be created by common mistyping errors such as missing a character, adding an additional character or transposing one or more characters within the target domain name.

The typo-neighborhood scanner takes the list of domains spawned by the typo-neighborhood generator and attempts to connect with each of them to determine if they exist and what sort of content they are serving. To prevent interference or issues from one typo-domain to the next, Strider uses a new Virtual Machine instance to connect with each one.

Using a modified version of the [Strider HoneyMonkey Scanner](#), the typo-neighborhood scanner uses a bank of 17 servers to execute the scans and obtain information about the typo-domains such as the third-party URLs visited and the content of all HTTP requests and responses. It can also be configured to capture a screen shot of the typo-domain site.

The Strider typo-domain database collects and analyzes the scan results. The data is then analyzed in three different ways. The typo-domain database looks at the typo-domains in a given category to determine how prevalent typo-squatting is for that category and who the culprits are behind the typo-squatting.

Secondly, the Strider typo-domain database examines the traffic to identify anchor domains. An anchor domain is a domain used to aggregate typo-squatting traffic from multiple typo-domains in order to simplify operations and revenue collection by nefarious website owners through one site. Determining the anchor domain provides a central point of reference for investigating and/or prosecuting typo-domain issues.

The third type of analysis is to search for specific key words, such as sexually explicit terms. The Strider typo-domain database reviews the HTTP response pages to extract typo-domains that contain any of the identified keywords.

Strider typo-neighborhood generator

The typo-neighborhood consists of all domains that are similar to, or potential typos of, the true target domain. The Strider typo-neighborhood generator uses five methods to generate the typo-domains that commonly occur:

- **Missing-dot typos:** These typos occur when a user fails to type the ".", or dot between the "www" and the domain name in the URL. For example, typing `http://wwwsecurityfocus.com` rather than `http://www.securityfocus.com`.
- **Character-omission typos:** These typo-domains are created by leaving out a letter of the domain name, one letter at a time. For example, `http://www.securityfocs.com` and `http://www.securityfous.com`.
- **Character-permutation typos:** These are domains that occur when two of the

letters in the domain name are transposed, or swapped while typing. Typo-neighborhood generator generates all such domains by swapping all characters one pair at a time. For example, `http://www.securiytfocus.com` or `http://www.securityfcous.com`.

- **Character-replacement typos:** To generate character-replacement typo-domains, the Strider typo-neighborhood generator replaces each letter in the domain with each of the letters adjacent to it on the keyboard. For example, typing `http://www.secueityfocus.com` or `http://www.securityfpcus.com`.
- **Character-insertion typos:** These typo domains are generated by inserting an additional character from one of the letters adjacent to the letter from the domain. It can also include using the same letter twice. For example, `http://www.securiotyfocus.com` or `http://www.securityffocus.com`.

Strider URL Tracer tool

The Microsoft research team created the [Strider URL Tracer](#) to work with and contribute to the typo-squatting project. The Strider URL Tracer performs four different functions to help users be aware of and have more control over traffic to third-party sites.

The URL Tracer tool has a URL Scan History function which logs and timestamps each primary URL that is visited, along with the third-party URLs that are communicated with as well. There is also a Top Domains view which displays the third-party domains that were visited, and lists each of the primary URLs visited that generated traffic to that third-party URL.

In both the URL Scan History and the Top Domains views, users can right-click on the domain names in the display and select Go or Block. Choosing Go will take the user to the URL to help identify which ads or traffic came from which URL's. Clicking on Block will prevent all future traffic to or from the identified domain.

Practical applications of the Strider URL Tracer with Typo-Patrol

Strider URL Tracer with Typo-Patrol can be used for two primary functions. One is to protect children from seeing inappropriate or explicit sites that they should not see and the other is for companies or trademark owners to scan and investigate sites that may be typo-squatting their domain(s) so that they can be investigated and / or prosecuted.

Protecting Children With Strider URL Tracer

If parents see that inappropriate or explicit ads are being displayed when their children are using the Web, they can refer to the URL Scan History to see what domains have been visited. Using the Go button, they can investigate which ad servers generated the illicit ads.

After the URL responsible for serving the unwanted ads is identified, parents can choose Block to ensure that no traffic goes to or from that site in the future and thereby protect their children from unscrupulous typo-squatters.

Protecting Trademarks With Strider URL Tracer

Trademark owners and companies can use the Strider URL Tracer to monitor their domain and identify typo-squatters and domain-parkers that may be infringing on or violating their trademark.

There are hundreds upon hundreds of possible typo-domains generated by using the Strider Typo-Neighborhood Generator. Most companies do not have the time, money or other resources required to try and investigate each of them to determine if there is any trademark violation and follow through with any kind of prosecution.

The Strider URL Tracer tool's Top Domains feature can be used to identify the largest, or most flagrant typo-squatters so that technical and legal resources can be invested where they will have the biggest effect.

The results can also be viewed grouped by IP address. Using these results, groups of typo-squatting domains that share an ISP can be identified and companies can send a multi-domain takedown notice to address a number of domains in one shot.

Avoid paying to advertise on a typo website

There is one other issue that arises from the systematic use of typo-squatting domains on domain-parking services and the use of syndicated web advertising. Sometimes a company may end up paying for a Web ad that is displayed to a user who was trying to go to the company web site in the first place.

For example, a user may want to visit `www.ford.com` and accidentally type `www.foed.com`. The user then ends up on a parked domain serving syndicated ad links, some of which are from Ford. So, even though the user actually meant to go straight to the Ford Web site, their typo redirected them to a different site which in turn charges Ford for the privilege of advertising to the user, who was effectively stolen from them in the first place.

Battling Typo-Squatters

According to the research and analysis done by the Microsoft Cybersecurity and Systems Management research group with the Strider Typo-Patrol project, between 40% and 70% of all active typo-squatting domains can be traced back to six major domain-parking services. One large-scale typo-squatter accounts for just under 20% of all the active typo-domains.

There are many laws that can be applied to fight typo-squatting. If a company's trademark or copyright has been infringed in any way, they can pursue damages for that. There are also laws aimed specifically at ensuring that frivolous domains are not set up, such as the

[Truth In Domain Names Act](#) or the [Anticybersquatting Consumer Protection Act](#).

Not every domain that happens to fit the typo-squatting algorithm is necessarily illegal, or even questionable, however. Some may be legitimate sites. One should use the Strider URL Tracer with Typo-Patrol to determine the extent of any typo-squatting related to your company domains or trademarks and identify the key players behind it - and then let the lawyers sort it out.

Conclusion

In this article, the author has discussed the need for Microsoft's free Strider URL Tracer with Typo-Patrol, to help fight typo-squatters and domain parking abuse. With numerous issues arising from typo-squatters and so many possible variations of such website names, the tools can be used by security professionals and home use alike.

References

- Microsoft research's free [Strider Typo-Patrol tool](#)
- The [Microsoft's .NET framework](#), required for Strider Typo-Patrol
- The [Strider HoneyMonkey Scanner](#), a project to detect and analyze Websites that may be hosting malicious code. See also <http://www.securityfocus.com/news/11273>
- [Truth In Domain Names Act](#)
- [Anticybersquatting Consumer Protection Act](#)

About the author

[Tony Bradley](#) is a [consultant and writer](#) who focuses on network security, antivirus and incident response. He is the About.com Guide author for Internet/Network Security, co-author of Hacker's Challenge 3, and author of the upcoming book Essential Computer Security. He also contributes frequently to other industry publications.

[Privacy Statement](#)

Copyright 2006, SecurityFocus