Hijacking Google search results for fun, not profit: UK SEO uncovers XML sitemap exploit in Google Search Console

In 2017, Google paid nearly $3 million to individuals and researchers as part of their Vulnerability Reward Program (VRP), which encourages the security research community to

find and report vulnerabilities in Google products.

This week, Tom Anthony – who heads Product Research & Development at Distilled, an SEO agency – was awarded a bug bounty of $1,337 for discovering an exploit that enabled one site to hijack the search engine results page (SERP) visibility and traffic of another – quickly getting indexed and easily ranking for the victimized site's competitive keywords.

Detailed in his blog post, Anthony describes how Google's Search Console (GSC) sitemap submission via ping URL essentially allowed him to submit an XML sitemap for a site he does control, as if it were a sitemap for one he does not. He did this by first finding a target site that allowed open redirects; scraping its contents and creating a duplicate of that site (and its URL structures) on a test server. He then submitted an XML sitemap to Google (hosted on the test server) that included URLs for the targeted domain with hreflang directives pointing to those same URLs, now also present on the test domain.

Hijacking the SERPs

Within 48 hours, the test domain started receiving traffic. Within the week, the test site was ranking for competitive terms on page 1 of the SERPs. Also, GSC showed the two sites as related – listing the targeted site as linking to the test site:

Google Search Console links the two unrelated sites. Source: http://www.tomanthony.co.uk

This presumed relationship also allowed Anthony to submit other XML sitemaps – within the test site's GSC at this point, not via ping URL – for the targeted site:

Victim site sitemap uploaded directly in GSC – Source: http://www.tomanthony.co.uk

Understanding the scope

Open redirects themselves are not a new or novel problem – and Google has been warning webmasters about shoring up their sites against this attack vector since 2009. What is noteworthy here is that utilizing an open redirect *worked* to not just submit a rogue sitemap, but to effectively rank a brand-new domain, brand-new site, with zero actual inbound links, and no promotion. And then to get that brand-new site and domain over a million search impressions, 10,000 unique visitors and 40,000 page views (via search traffic only) in

three weeks.

The “bug” here is both a problem with sitemap submissions (the subsequent sail-through GSC sitemap submissions are alarming) and a greater problem as to how the algorithm immediately applied all the equity from the one site across to the completely separate and unrelated domain.

Source: http://www.tomanthony.co.uk

I reached out to Google with a series of detailed questions about this exploit, including the search quality team’s involvement in pursuing and implementing a fix, and whether or not they are able to detect and take action on any bad actors that may have already exploited this vulnerability. A Google

spokesperson replied:

> *When we were alerted to the issue, we worked closely across teams to address it. It was not a previously known issue and we don't believe it had been used.*

In response to questions about changes with respect to sitemap submissions, GSC and the transfer of equity affecting results, the spokesperson said:

> *We continue to recommend that site-owners use sitemaps to let us know about new & updated pages within their website. Additionally, the new Search Console also uses sitemaps as a way of drilling down into specific information within your website in the Index Coverage report. If you're hosting your sitemaps outside of your website, for proper usage it's important that you have both sites verified in the same Search Console account.*

I discussed this exploit and the research at length with Anthony.

The research process

When asked about his motivations for pursuing this work, he said, "I believe an effective SEO is someone who experiments and tries to understand things behind the scenes. I've never done any black-hat SEO, and so set myself the challenge of finding something on that side of things; primarily for the learning experience and as a way to run defense if I ever saw it in the wild."

He added, "I like doing security research as a hobby on the side, so decided that rather than take the 'traditional' black-hat route of manipulating the algorithm's ranking signals, I'd see if I could instead find an outright bug in it."

Oftentimes, the driving motivation in pursuing a given method relates to having experienced (or having a client that has experienced) a sudden drop in SERP traffic or rankings. Anthony noted, “At Distilled, like so many SEOs, I’ve worked with sites that have had unexplained drops. Often clients claim ‘negative SEO,’ but usually it is something far more mundane. What is worrying about this specific issue is [that] typical negative SEO attacks are detectable. If I spam you with low-quality links, you can find them, you can confirm they exist. With this issue, it appears an attacker could leverage your equity in Google and you would not know.”

Over the course of four weeks’ evenings and weekends spent delving into it, Anthony discovered that combining different research streams he’d begun proved effective where each separately led to dead ends. “I had ended up with two threads of research – one around open redirects, as they are a crack in how sites work that I felt could be leveraged for SEO – and the other was with XML sitemaps and trying to make Googlebot error out when parsing them (I ran about 20 variations of that, but none worked!). I was so deep into it at this point, and had a revelation when I realized these two streams of research could perhaps be combined.”

Reporting and resolution

Once he realized the impact and harm that could be done to sites, Anthony reported the bug to Google’s security team (See complete timeline in his post). As this method was previously unknown to Google but clearly exploitable, Anthony noted, “It is a terrifying prospect that this could have already been out there and being exploited. However, the nature of the bug would mean it is essentially undetectable. The ‘victim’ may not be affected directly if their equity is used to rank in another country, and then the victims become the legitimate companies who are pushed down the rankings by the attacker. They would have no way to tell how the attacker site was

ranking so well.”

As noted above, the Google spokesperson said they do not believe it has been used. Unclear from their response is whether or not they have data available that would enable them to detect pinged sitemaps used in such a way. If further comment or information is given, we’ll update this post.

On the issue of detection specifically, I asked Anthony to speculate on scaling this exploit. “The biggest weakness with my experiment was how closely I mimicked the original site in terms of URL structure and content. I had a bunch of experiments prepped that were designed to measure just how different you could make the attacker site: Do I need the same URL structure as the parent site? How similar must the content be? Can I target other languages in the same country as the victim site? In my case, I think I could have re-run with the same approach but have differentiated the attack site slightly more, and probably [would] have escaped detection,” he said.

He added, “If I had kept it to myself, then I imagine I could have gone for months or years. If you outright scammed people it would be short-lived, but if you used the method to drive affiliate traffic, or even simply to boost your own legitimate business, then little reason you’d ever be caught.”

As the image below demonstrates, the short-lived traffic driven to the test site was potentially far more valuable than the relatively small (by comparison) bounty he was awarded, which makes one wonder if the security team really understood the implications of the exploit.

Searchmetrics' Traffic Value. Source: http://www.tomanthony.co.uk

Anthony's motivations (and why he did report the vulnerability right away) were rooted in research and helping the search community, however.

"Doing this sort of research is a learning experience, and not about abusing what you find. In the industry, we have our complaints about Google at times, but [to] a consumer, they provide a great service, and I think good SEOs actually help with that – and this is basically an extension of the same idea. The Vulnerability Reward Program they run is a nice incentive to focus research efforts on them rather than elsewhere; it is nice to potentially be awarded a bounty for

the time and effort that goes into the research.”

About The Author

Michelle Robbins, SVP Content & Marketing Technology, oversees editorial direction as Editor in Chief for Third Door Media’s digital publications, Search Engine Land, Marketing Land and MarTech Today, directing a full-time staff of reporters and editors managing contributed content. She is responsible for developing the content strategy across all properties and aligning those initiatives with the programming and audience goals for Third Door Media’s two leading marketing conference series, Search Marketing Expo and The MarTech Conference. In addition, Michelle oversees information technology operations, directing the marketing technology department. An experienced domestic and international keynote and featured speaker, she enjoys connecting with the community at SMX, MarTech and other industry events. Connect online with Michelle on Twitter @MichelleRobbins, and Linkedin.
http://feeds.searchengineland.com/~r/searchengineland/~3/uGLr7J6JVd8/hijacking-google-search-results-for-fun-not-profit-uk-seo-uncovers-xml-sitemap-exploit-in-google-search-console-295169