Remedying Web Hijacking: Notification Effectiveness and Webmaster Comprehension

Frank Li[†] Grant Ho[†] Eric Kuan[◇] Yuan Niu[◇]
Lucas Ballard[◇] Kurt Thomas[◇] Elie Bursztein[◇] Vern Paxson[†*]

{frankli, grantho, vern}@cs.berkeley.edu {erickuan, niu, lucasballard, kurtthomas, elieb}@google.com

[†]University of California, Berkeley [◇]Google Inc. [*]International Computer Science Institute

ABSTRACT

As miscreants routinely hijack thousands of vulnerable web servers weekly for cheap hosting and traffic acquisition, security services have turned to notifications both to alert webmasters of ongoing incidents as well as to expedite recovery. In this work we present the first large-scale measurement study on the effectiveness of combinations of browser, search, and direct webmaster notifications at reducing the duration a site remains compromised. Our study captures the life cycle of 760,935 hijacking incidents from July, 2014–June, 2015, as identified by Google Safe Browsing and Search Quality. We observe that direct communication with webmasters increases the likelihood of cleanup by over 50% and reduces infection lengths by at least 62%. Absent this open channel for communication, we find browser interstitials—while intended to alert visitors to potentially harmful content—correlate with faster remediation. As part of our study, we also explore whether webmasters exhibit the necessary technical expertise to address hijacking incidents. Based on appeal logs where webmasters alert Google that their (site)-265is no longer compromised, we find 80% of operators successfully clean up symptoms on their first appeal. However, a

sites (as captured by search ranking) are three times more likely to clean up in 14 days compared to unpopular sites. Our results illustrate that webmasters benefit significantly from detailed, external security alerts, but that major gaps exist between the capabilities of small websites and major institutions.

To better understand the impact of webmaster comprehension on overall remediation, we decouple the period before webmasters receive a notification from the time spent cleaning a site. We capture this behavior based on appeals logs that detail when webmasters request Google to remove hijacking warnings from their site, a request verified by a human analyst or crawler. We find that 80% of site operators successfully clean up on their first appeal attempt. The remaining 20% require multiple appeals, spending a median of one week purging attacker code. Equally problematic, many site operators appear to address only symptoms rather than the root cause of compromise: 12% of sites fall victim to hijacking again within 30 days, with 10% and 20% of re-infections occurring within one day for Safe Browsing and Search Quality respectively. We distill these pain points into a path forward for improving remediation. In the process, we highlight the tension between protecting users and webmasters and the decentralized responsibilities of Internet security that ultimately confound recovery.

In summary, we frame our key contributions as follows:

- We present the first large-scale study on the impact of diverse notification strategies on the outcome of 760,935 hijacking incidents.
- We model infection duration, finding that notification techniques outweigh site popularity or webmaster knowledge as the most influential factor for expediting recovery.
- We find that some webmasters struggle with the remediation process, with over 12% of sites falling victim to re-compromise in 30 days.
- We present a path forward for helping webmasters recover and the pitfalls therein.

2. BACKGROUND & RELATED WORK

Web services rely on *notifications* to alert site operators to security events after a breach occurs. In this study, we focus specifically on website compromise, but notifications extend to account hijacking and credit card fraud among other abuse. We distinguish notifications from *warnings* where visitors are directed away from unsafe sites or decisions (e.g., installing an unsigned binary). With warnings, there is a path back to safety and minimal technical expertise is required; breaches lack this luxury and require a more complex remedy that can only be addressed by site operators. We outline approaches for identifying compromise and prior evaluations on notification effectiveness.

2.1 Detecting Compromise

As a precursor to notification, web services must first detect compromise. The dominant research strategy has been to identify side-effects injected by an attacker. For example, Provos et al. detected hacked websites serving drive-by downloads based on spawned processes [18]; Wang et al. detected suspicious redirects introduced by blackhat cloaking [26]; and Borgolte detected common symptoms of defacement [3]. These same strategies extend beyond the web arena to detecting account hijacking, where prior work relied on identifying anomalous usage patterns or wide-scale collusion [7, 22]. More recently, Vasek et al. examined factors that influenced the likelihood of compromise, including the serving platform (e.g., nginx) and content management system (e.g., Wordpress) [25]. They found that sites operating popular platforms such as Wordpress, Joomla, and Drupal faced an increased risk of becoming compromised, primarily because miscreants focused their efforts on exploits that impacted the largest market share. Soska et al. extended this idea by clustering websites running the same version of content management systems in order to predict likely outbreaks of compromise [19]. We sidestep the issue of detection, instead relying on a feed of known compromised pages involved in drive-bys, spam, or cloaking (§ 3).

2.2 Webmaster Perspective of Compromise

Given a wealth of techniques to detect compromise, the preeminent challenge that remains is how best to alert webmasters to security breaches, assuming webmasters are incapable of running detection locally. StopBadware and CommTouch surveyed over 600 webmasters of compromised websites to understand their process for detecting compromise and remedying infection [21]. They found that only 6% of webmasters discovered an infection via proactive monitoring for suspicious activity. In contrast, 49% of webmasters learned about the compromise when they received a browser warning while attempting to view their own site; another 35% found out through other third-party reporting channels, such as contact from their web hosting provider or a notification from a colleague or friend who received a browser warning.

Equally problematic, webmasters rarely receive support from their web hosting providers. Canali et al. created vulnerable websites on 22 hosting providers and ran a series of five attacks that simulated infections on each of these websites over 25 days [4]. Within that 25-day window, they found that only one hosting provider contacted them about a potential compromise of their website, even though the infections they induced were detectable by free, publicly-available tools. Similarly, the StopBadware and CommTouch study found that 46% of site operators cleaned up infections themselves, while another 20% reached out for professional help [21]. Only 34% of webmasters had the option of free help from their hosting provider (irrespective of using it). These two studies provide qualitative evidence of the struggles currently facing webmasters and the potential value of third-party notifications.

2.3 Measuring the Impact of Notifications

A multitude of studies previously explored the impact of notifications on the likelihood and time frame of remediation. Vasek et al. examined the impact of sending malware reports to 161 infected websites [24]. They emailed each site's owner (either via manual identification or WHOIS information) and found 32% of sites cleaned up within a day of notification, compared to 13% of sites that were not notified. Cleanup was further improved by providing

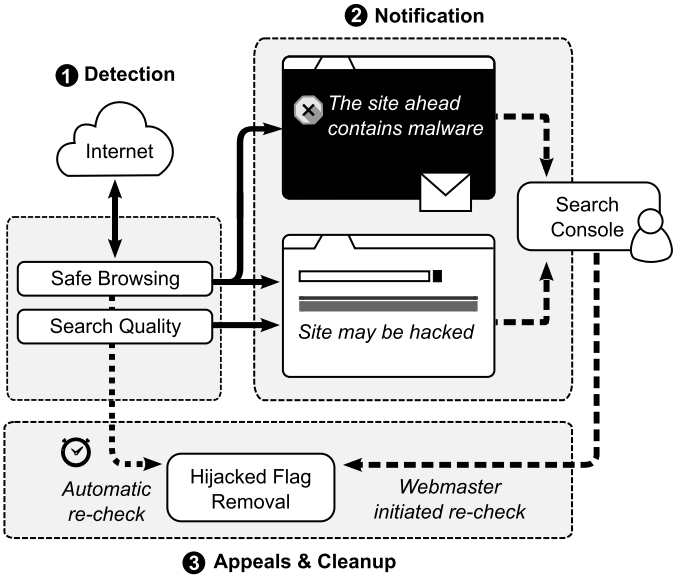


Figure 1: Google's hijacking notification systems. Safe Browsing and Search Quality each detect and flag hijacked websites (❶). From there, Safe Browsing shows a browser interstitial and emails WHOIS admins, while both Safe Browsing and Search Quality flag URLs in Google Search with a warning message (❷). Additionally, if the site's owner has a Search Console account, a direct notification alerts the webmaster of the ongoing incident. The hijacking flag is removed if an automatic re-check determines the site is no longer compromised. Webmasters can also manually trigger this re-check via Search Console (❸).

to 26.8% of unnotified operators. Similarly, Kührer et al. identified over 9 million NTP amplifiers that attackers could exploit for DDoS and relied on a public advisory via MITRE, CERT-CC, and PSIRT to reach administrators. After one year, they found the number of amplifiers decreased by 33.9%, though lacked a control to understand the precise impact of notification. Each of these studies established that notifications decrease the duration of infections. In our work, we explore a critical next step: whether combinations of notifications and warnings reach a wider audience and ultimately improve remediation.

3. METHODOLOGY

Our study builds on data gathered from two unique production pipelines that detect and notify webmasters about compromise: Safe Browsing, which covers drive-by attacks [18]; and Search Quality, which protects against scams and cloaked gateways that might otherwise pollute Google Search [9]. We describe each pipeline's detection technique, the subsequent notification signals sent, and how each system determines when webmasters clean up. A high level description of this process appears in Figure 1. We note that our measurements rely on *in situ* data collection; we cannot modify the system in any way, requiring that we thoroughly account for any potential biases or limitations. We provide a breakdown of our final dataset in Table 1, collected over a time frame of 11 months from July 15th, 2014–June 1st, 2015.

3.1 Compromised Sites

When Safe Browsing or Search Quality detect a page hosting harmful or scam content, they set a flag that subsequently triggers both notifications and warnings. We group flags on the level of registered domains (e.g., *example.com*), with the exception of shared web hosting sites, where we consider flags operating on subdomains (e.g., *example.blogspot.com*). Both Safe Browsing and Search Quality distinguish purely malicious pages from compromised sites based on whether a site previously hosted legitimate content; we restrict all analysis to compromised sites. For the purposes of our study, we denote a *hijacked website* as any registered or shared domain that miscreants compromise. We use the term *hijacking incident* to qualify an individual attack: if miscreants subvert multiple pages on a website, we treat it as a single incident up until Safe Browsing or Search Quality verify the site is cleaned (discussed in § 3.3).[1] We treat any subsequent appearance of malicious content as a new hijacking incident.

Dataset	Safe Browsing	Search Quality
Time frame	7/15/14–6/1/15	7/15/14–6/1/15
Hijacked websites	313,190	266,742
Hijacking incidents	336,122	424,813
Search console alerts	51,426	88,392
WHOIS emails	336,122	0
Webmaster appeals	124,370	48,262

Table 1: Summary of dataset used to evaluate notification effectiveness and webmaster comprehension.

As detailed in Table 1, we observe a total of 760,935 such incidents, with the weekly breakdown of new incidents shown in Figure 2. Our dataset demonstrates that miscreants routinely compromise new websites, with a median of 8,987 new sites detected by Search Quality and 5,802 sites by Safe Browsing each week. We also find evidence of rare, large-scale outbreaks of compromise that impact over 30,000 sites simultaneously.

We caution that our dataset is biased to hijacking threats known to Google and is by no means exhaustive. From periodic manual reviews performed by Google analysts of a random sample of hijacking incidents, we estimate the false positive rates of both pipelines to be near zero, though false negatives remain unknown. That said, our dataset arguably provides a representative cross-section of hijacked webpages around the globe. We provide a detailed demographic breakdown later in § 4.

3.2 Notification Mechanisms

Safe Browsing and Search Quality each rely on a distinct combination of notification and warning mechanisms to alert webmasters either directly or indirectly of hijacking incidents. We detail each of the possible combinations in Figure 1, spanning browser interstitials, search page warnings, webmaster help tools (called Search Console), and WHOIS emails. For all of these notification mechanisms, we lack visibility into whether webmasters observe the alert (e.g., received and read a message or viewed a browser interstitial). As such, when we measure the effectiveness of notifications, we couple both the distribution of the signal and the subsequent webmaster response.

Browser Interstitials: Safe Browsing reports all compromised websites to Chrome, Safari, and Firefox users that opt for security warnings. While browser interstitials primarily serve to warn visitors of drive-by pages that cause harm, they also serve as an indirect alerting mechanism to site owners: webmasters (or their

[1] In the event multiple attackers compromise the same site simultaneously, we will mistakenly conflate the symptoms as a single hijacking incident.

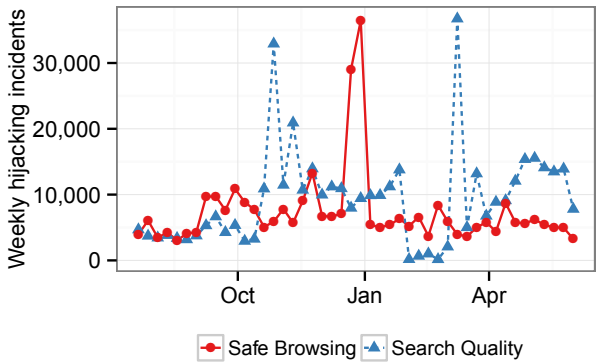


Figure 2: Weekly breakdown of new website hijacking incidents as detected by Safe Browsing (drive-bys) and Search Quality (scams, blackhat SEO, cloaking).

peers) that encounter warnings for their compromised sites can take action. However, this approach lacks diagnostic information about how exactly to clean up the infection.

Search Result Annotation: Compromised sites detected by Search Quality receive an annotation in search result pages with a warning "This site may be hacked" [13]. Similarly, sites identified by Safe Browsing all receive a flag "This site may harm your computer" [14]. Additionally, sites may lose search ranking. As with browser interstitials, these modifications primarily serve to protect inbound visitors, but also serve as an indirect channel for alerting webmasters of compromise.

Search Console Alerts: Both Safe Browsing and Search Quality provide concrete details about infected pages, examples, and tips for remediation via Google's Search Console [12]. Only webmasters who register with Search Console can view this information. Notifications come in the form of alerts on Search Console as well as a single message to the webmaster's personal email address (provided during registration). This approach alleviates some of the uncertainty of whether contacting the WHOIS `abuse@` will reach the site operator. One caveat is that a Search Console notification is sent only if the Safe Browsing or Search Quality algorithm is configured to do so.[2] This behavior results in the absence of a Search Console alert for some hijacking incidents, which we use as a natural control to measure their effectiveness.

Of hijacking incidents identified by Search Quality, 95,095 (22%) were sites where webmasters had registered with Search Console prior to infection. The same is true for 107,581 (32%) of Safe Browsing hijacking incidents. Of these, 48% of incidents flagged by Safe Browsing triggered a Search Console alert, as did 93% of incidents identified by Search Quality. Because notified and unnotified sites are flagged by different detection subsystems, differences in the subsystems may result in a biased control population. However, the bias is likely modest since the differences in detection approaches are fairly subtle compared to the overall nature of the compromise.

WHOIS Email: In an attempt to contact a wider audience, Safe Browsing also emails the WHOIS admin associated with each compromised site. Since Safe Browsing simultaneously displays warnings via browser interstitials and search, we can only measure the aggregate impact of both techniques on notification effectiveness.

[2]Each detection pipeline combines a multitude of algorithms, of which the majority also generate Search Console alerts.

3.3 Appeals & Cleanup

Safe Browsing and Search Quality automatically detect when websites clean up by periodically scanning pages for symptoms of compromise until they are no longer present. This approach is limited by the re-scan rate of both pipelines. For Safe Browsing, sites are eligible for re-scans 14 days after their previous scan. This metric provides an accurate signal on the eventual fate of a hijacking incident, but leaves a coarse window of about 14 days during which a site may clean up but will not be immediately flagged as such. Search Quality reassesses symptoms each time Google's crawler revisits the page, enabling much finer-grained analysis windows.

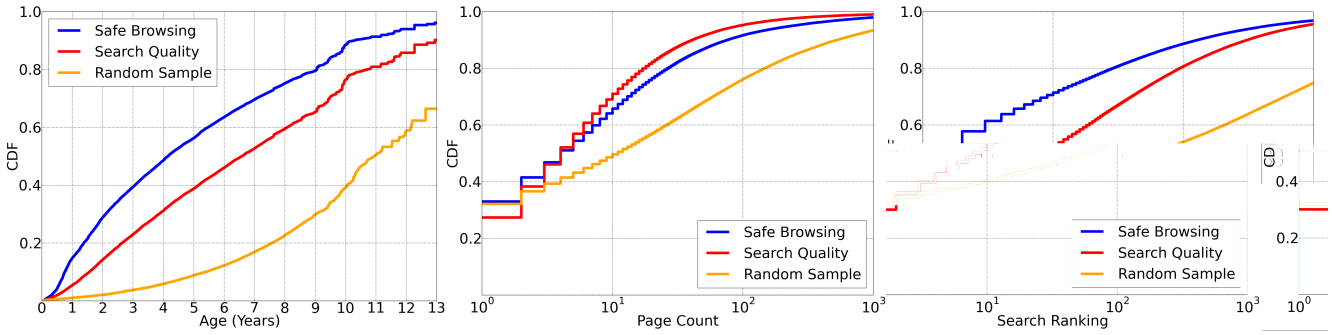
As an alternative, both Safe Browsing and Search Quality provide a tool via Search Console where webmasters can appeal warnings tied to their site. Safe Browsing webmasters have an additional channel of appealing through StopBadware [20], which submits requests for site review to Safe Browsing on behalf of the webmasters who chose not to register with Search Console. The appeals process signals when webmasters believe their pages are cleaned, which analysts at Google (Search Quality) or automated re-scans (Safe Browsing) subsequently confirm or reject. We use this dataset both to measure the duration of compromise as well as the capability of webmasters to clean up effectively, as captured by repeatedly rejected appeals.

3.4 Limitations

We highlight and reiterate five limitations tied to our dataset. First, as a natural experiment our study lacks a true control: the systems under analysis notify all webmasters, where Google's policies arbitrate the extent of the notification. For example, browser interstitials are only shown for threats that may put site visitors at risk, namely Safe Browsing sites. Webmaster notifications via Search Console are sent only when a webmaster has registered for Search Console and the specific detection pipeline is integrated with the Search Console messaging platform. As such, we restrict our study to comparing the relative effectiveness of various notification approaches; previous studies have already demonstrated the value of notifications over a control [5,6,24]. Second, our coverage of hijacked websites is biased towards threats caught by Google's pipelines, though we still capture a sample size of 760,935 incidents, the largest studied to date. Third, when victims are notified, we lack visibility into whether the intended recipient witnesses the alerts. As such, when we measure the impact of notifications, we are measuring both the distribution of alerts and the ability of webmasters to take action. Fourth, the granularity at which we measure cleanup is not always per-day, but may span multiple days before a site is rechecked. This may cause us to overestimate the time a site remains compromised. We specifically evaluate compromise incidents as continuous blacklisting intervals. Finally, our study is not universally reproducible as we rely on a proprietary dataset. However, given the scale and breadth of compromised websites and notifications we analyze, we believe the insights gained are generalizable.

4. WEBSITE DEMOGRAPHICS

As a precursor to our study, we provide a demographic breakdown of websites (e.g., domains) that fall victim to hijacking and compare it against a random sample of 100,000 non-spam sites indexed by Google Search. All of the features we explore originate from Googlebot, Google's search crawler [10]. We find that compromise affects webpages of all age, language, and search ranking, with attackers disproportionately breaching small websites over major institutions. We also observe biases in the sites that miscreants re-purpose for exploits versus scams and cloaked gateways.



(a) Site Age **(b)** Page Count **(c)** Search Rankings

Figure 3: Demographic breakdowns of sites flagged as hijacked by Safe Browsing and Search Quality.

Site Age: We estimate the age of a site as the time between Googlebot's first visit and the conclusion of our study in June, 2015. We detail our results in Figure 3a. We find over 85% of Search Quality sites are at least two years old at the time of compromise, compared to 71% of Safe Browsing sites. Our results indicate that sites falling victim to compromise skew towards newer properties compared to the general population, more so for Safe Browsing.

Page Count: We estimate a site's size as the total number of pages per domain indexed by Googlebot at the conclusion of our study. In the event a site was hijacked, we restrict this calculation to the number of pages at the onset of compromise so as not to conflate a proliferation of scam offers with prior legitimate content. Figure 3b shows a breakdown of pages per domain. Overall, if we consider the volume of content as a proxy metric of complexity, we observe that hijacking skews towards smaller, simpler sites compared to larger properties run by major institutions.

Search Page Rankings: The search result ranking of a site represents a site's popularity, as captured by a multitude of factors including page rank and query relevance. We plot the distribution of search ranking among sites in Figure 3c. As an alternative metric to search ranking, we also calculate the Alexa ranking of hijacked sites. We observe that Search Quality sites skew towards higher search rankings compared to Safe Browsing, a bias introduced by cloaking and scam-based attacks targeting sites more likely to draw in visitors from Google Search. There are limits, however: we find less than 5% of compromised properties appear in the Alexa Top Million. Similarly, compared to our random sample, hijacking disproportionately impacts lowly-ranked pages. Overall, 30% of Search Quality sites and 50% of Safe Browsing sites rank low enough to receive limited search traction. This suggests that search result warnings may be ineffective for such properties due to limited visibility, a factor we explore later in § 5.

Language: For each site, we obtain Googlebot's estimate of the site's primary language, shown in Table 2. We find that miscreants predominantly target English, Russian, and Chinese sites, but all languages are adversely affected. We observe a substantially different distribution of languages than the work by Provos et al., which studied exclusively malicious properties [18]. Their work found that miscreants served over 60% of exploits from Chinese sites and 15% from the United States. This discrepancy re-iterates that exploit delivery is often a separate function from web compromise, where the latter serves only as a tool for traffic generation. We observe other examples of this specialization in our dataset: 10% of Safe Browsing incidents affect Chinese sites, compared to only 1% of Search Quality incidents. Conversely, Search Quality is far more likely to target English sites. These discrepancies hint at potential regional biases introduced by the actors behind attacks.

Language	Rnd. Sample	Safe Browsing	Search Quality
English	35.4%	46.9%	57.6%
Chinese	9.0%	10.0%	1.0%
German	7.2%	4.5%	2.5%
Japanese	6.0%	1.5%	4.6%
Russian	5.6%	9.9%	6.3%

Table 2: Top five languages for a random sample of websites versus sites falling victim to compromise. We find hijacking disproportionately impacts English, Russian, and Chinese sites.

Software	Rnd. Sample	Safe Browsing	Search Quality
WordPress	47.4%	36.9%	39.6%
Joomla	10.7%	11.6%	20.4%
Drupal	8.3%	1.7%	9.4%
Typo3	3.4%	0.5%	0.9%
Vbulletin	3.3 %	0.8%	0.4%
Discuz	1.0%	7.6%	0.4%
DedeCMS	0.2%	13.9%	1.4%

Table 3: Top five software platforms for hijacked websites and significant outliers, when detected. We find attacks target all of the top platforms. We note these annotations exist for only 10–13.8% of sites per data source.

Site Software: Our final site annotation consists of the content management system (CMS) or forum software that webmasters serve content from, when applicable. Of hijacked sites, 9.1% from Safe Browsing and 13.8% from Search Quality include this annotation, compared to 10.3% of sampled sites. We present a breakdown of this subsample in Table 3. Wordpress and Joomla are both popular targets of hijacking, though this partly represents each system's market share. The popularity of DedeCMS and Discuz among Safe Browsing sites might be unexpected, but as Chinese platforms, this accords with the large proportion of Chinese Safe Browsing sites. Studies have shown that historically attackers prioritize attacking popular CMS platforms [25].

5. NOTIFICATION EFFECTIVENESS

We evaluate the effect of browser warnings, search warnings, and direct communication with webmasters on remediation along two dimensions: the overall duration a site remains compromised and the fraction of sites that never clean up, despite best efforts to alert the respective webmaster. We also consider other influential factors such as whether symptoms of compromise are localized to a single page or systemic to an entire site; and whether language barriers impede effective notification. To avoid bias from webmasters with prior exposure to infection, we restrict our analysis to the first incident per site.[3] (We explore repeated hijacking incidents later in § 6.) We reiterate that our experiments lack a control, as outlined in § 3, due to our *in situ* measurements of a live system that always sends at least some form of notification. As such, we restrict our evaluation to a comparison between notification approaches and contrast our findings with prior work.

5.1 Aggregate Remediation Outcomes

To start, we explore in aggregate the infection duration of all websites identified by Safe Browsing and Search Quality. We find that over our 11-month period, webmasters resolved 59.5% of hijacking incidents. Breaking this down into time windows, 6.6% of sites cleaned up within a day of detection, 27.9% within two weeks, and 41.2% within one month. Note that for the single-day cleanup rate, we exclude Safe Browsing sites that were automatically rescanned and de-listed as we have no remediation information until two weeks after an infection begins. This does not impact manually appealed Safe Browsing incidents or any Search Quality incidents. The 40.5% of sites that remain infected at the end of our collection window have languished in that state for a median of four months, with 10% of persistent infections dating back over eight months. Our results indicate a slower remediation rate than observed by Vasek et al., who found 45% of 45 unnotified hijacked sites and 63% of 53 notified sites cleaned up within 16 days [24].

We recognize these numbers may be biased due webmasters not having enough time to redress infections occurring towards the tail end of our collection window. If we repeat the aggregate analysis only for hijacking incidents with an onset prior to January 1, 2015 (the first half of our dataset), we find the resulting cleanup rate is 7% higher, or 66.7%. Breaking this down into time windows, we find only a slight difference compared to our earlier analysis of the entire dataset: 6.4% of sites cleaned up within a day after detection, 28.2% within two weeks, and 42.1% within one month. As such, given an infection that is older than a few months, we find the likelihood of cleanup in the future tends towards zero.

5.2 Ranking Notification Approaches

We find that notification approaches significantly impact the likelihood of remediation. Browser warnings, in conjunction with a WHOIS email and search warnings, result in 54.6% of sites cleaning up, compared to 43.4% of sites only flagged in Google Search as "hacked". For webmasters that previously registered with Search Console and received a direct alert, the likelihood of remediation increases to 82.4% for Safe Browsing and 76.8% for Search Quality. These results strongly indicate that having an open channel to webmasters is critical for overall remediation. Furthermore, it appears that browser interstitials (in conjunction with possible WHOIS email contact) outperform exclusive search warnings. Assuming the two compromise classes do not differ significantly

[3]We filter sites who have appeared on the Safe Browsing or Search Quality blacklists prior to our collection window. However, note it may be possible that sites were previously compromised and undetected. This filtering is best effort.

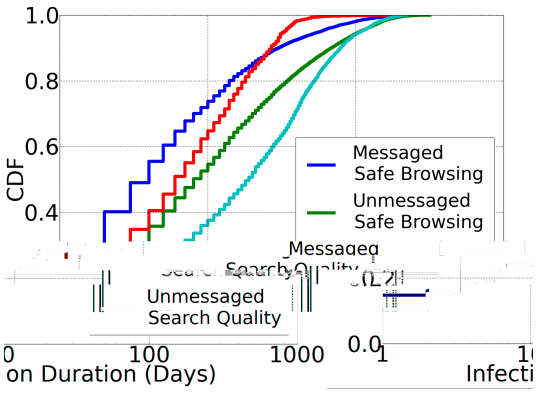


Figure 4: CDFs of the infection duration for Safe Browsing and Search Quality sites that eventually recover, dependent on whether they received Search Console messages or not. Direct communication with webmasters significantly reduces the duration of compromise.

in remediation difficulty, this suggests that browser warnings and WHOIS emails stand a better chance at reaching webmasters. This is consistent with our observation in § 4 that the majority of hijacked sites have a low search rank, impeding visibility of search warnings. Another possible explanation is that browser interstitials create a stronger incentive to clean up as Safe Browsing outright blocks visitors, while Search only presents an advisory.

For those sites that do clean up, we present a CDF of infection duration in Figure 4 split by the notifications sent. We omit the 37.8% sites that Safe Browsing automatically confirmed as cleaned from this analysis as we lack data points other than at a 14-day granularity. We note that after 14 days, 8% more manually appealed sites were cleaned compared to automatically appealed sites, indicating the latter population of site operators clean up at a slightly slower rate. As with remediation likelihood, we also observe that direct communication with webmasters decreases the time attackers retain access to a site. Within 3 days, 50% of Safe Browsing manually-appealed sites clean up when notified via Search Console, compared to 8 days in the absence of Search Console alerts. We observe a similar impact for Search Quality: 50% of webmasters resolve in 7 days, versus 18 days for unassisted sites. This indicates that incorporating direct, informative messages expedites recovery.

5.3 Localized vs. Systemic Infections

Given the positive influence of notifications, an important question remains whether the complexity of an infection influences the time necessary to clean up. While Search Console provides remediation tips, webmasters may naturally require more support (and detailed reports) to contend with systemic infections. To assess this, we rely on detailed logs provided by Search Quality on whether harmful content for hijacked properties was systemic to an entire site or localized. In particular, Search Quality attempts to categorize incidents into three groups: *isolated incidents*, where harmful content appears in a single directory (N=91,472); *dispersed incidents* that affect multiple URL paths and subdomains (N=21,945); and *redirection incidents*, where sites become a gateway to other harmful landing pages (N=21,627).

Figure 5 provides a breakdown of infection duration by hijacking symptoms irrespective of whether webmasters received a Search Console alert. This measurement also includes sites that have yet to recover. We find that webmasters are both more likely and quicker

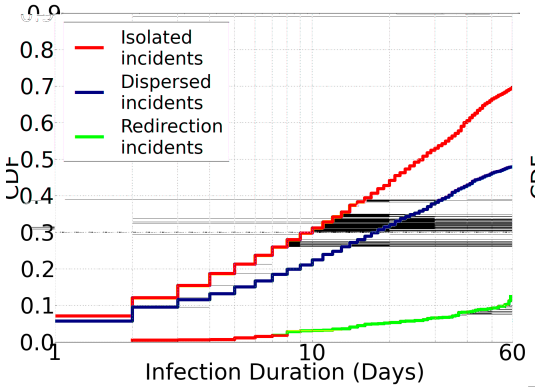


Figure 5: CDFs of the infection duration for different Search Quality incidents. We note curves do not reach 100% because we consider the percentage over all sites, including those not yet clean after 60 days. Cloaking (redirection) and systemic infections (dispersed) prove the most difficult for webmasters to contend with.

Language	Search Quality	Safe Browsing
English	24.7%	29.5%
Chinese	20.1%	2.7%
Russian	19.4%	22.3%
Spanish	24.6%	30.4%
German	22.9%	24.6%
French	20.9%	26.9%
Italian	25.8%	29.1%
Polish	26.4%	28.9%
Portuguese (Brazil)	24.1%	27.6%
Japanese	26.1%	28.2%

Table 4: Top 10 languages and remediation rates for Safe Browsing and Search Quality, restricted to sites not registered with Search Console. We find consistent rates despite languages differences, with the exception of Chinese due to low Safe Browsing usage in the region.

to clean up isolated incidents. Of all resolved and ongoing isolated incidents, only 50% last longer than 27 days. In contrast, over 50% of dispersed incidents persist for longer than 60 days. The most challenging incident type relates to redirect attacks where sites become cloaked gateways. Of these, only 12% recover within 60 days. One possible explanation for low cleanup behavior is that redirect-based attacks cloak against site operators, preventing webmasters from triggering the behavior [26]. Alternatively, redirection attacks require far less code (e.g., .htaccess modification, JavaScript snippet, meta-header) compared to attackers hosting entirely new pages and directories. As such, it becomes more difficult for webmasters to detect the modifications. In aggregate, our findings demonstrate that webmasters impacted by systemic or redirect incidents require far more effort to recover. We discuss potential aids later in § 7.

5.4 Notification Language

To avoid language barriers and assist victims of a global epidemic, Search Console supports over 25 popular languages [1]. We investigate the impact that language translations might have on recovery speeds, examining sites that received a Search Console message. Breaking down these hijacking incidents by site language reveals only relatively small differences in the fraction of sites cleaned after two weeks for the ten most popular languages. In particular, remediation rates vary between 10% for Safe Browsing and 7% for Search Quality, indicating that the message language does not drastically influence cleanup.

To evaluate the influence of language for the other notification vectors, we measure the cleanup rate for non-Search Console registrants. Table 4 lists the percentage of Search Quality and Safe Browsing sites recovered after two weeks. For Search Quality, remediation rates range 7%. Excluding Chinese sites, Safe Browsing sites are also similar, varying between 9%. Thus, language or geographic biases do not play a major factor for browser interstitials and search warnings, with the exception of Chinese sites.

Given Chinese sites present an outlier, we explore possible explanations. Despite having the largest Internet population in the world, China ranks 45th in the number of daily browser interstitials shown, a discrepancy unseen for countries who speak the other top languages. Since we find a significant number of Chinese sites serving malware, we argue this discrepancy is due to lagging adoption of Safe Browsing in China, and potentially explains the slow Chinese remediation rate for Safe Browsing. However, it remains unclear why Chinese Search Quality sites are comparable to other languages. As there are popular search alternatives local to China, Google's search result warnings may have low visibility. Alternatively, it may be possible that Chinese sites flagged by Search Quality are a different population that those for Safe Browsing; for example, these may be Chinese language sites with a target audience outside of China and can benefit from external signals.

5.5 Site Popularity

We correlate a site's popularity (e.g., search ranking) with 14-day remediation rates for hijacking incidents. In particular, we sort sites based on search ranking and then chunk the distribution into bins of at least size 100, averaging the search ranking per bin and calculating the median cleanup time per bin. We present our results in Figure 6. We find more popular sites recover faster across both Safe Browsing and Search Quality. Multiple factors may influence this outcome. First, popular sites have a strong incentives to maintain site reputation and business. Additionally, with more visitors and a higher search ranking, it is more likely that a site visitor encounters a browser interstitial or search page warning and informs the webmaster. We find this reaches a limit, with Safe Browsing and Search Quality cleanup rates converging for search rankings greater than one million. Finally, site popularity may also reflect companies with significantly higher resources and more technical administrators. Regardless the explanation, it is clear that less popular websites suffer from significantly longer infections compared to major institutions. As discussed in § 4, these lowly ranked sites comprise the brunt of hijacked properties.

5.6 Search Console Awareness

Webmasters who proactively register with Search Console may represent a biased, more savvy population compared to all webmasters. As such, we may conflate the improved performance of Search Console alerts with confounding variables. As detailed in § 3, not all hijacking incidents trigger a Search Console alert, even if the site owner previously registered with the service. We compare the likelihood of remediation for this subpopulation against the set of users who never register with Search Console. After two weeks, 24.5% of sites registered with Search Console and identified by Search Quality cleaned up, compared to 21.0% of non-registrants. Similarly for Safe Browsing, 28.4% of Search Console sites cleaned up compared to 24.3% of non-registrants. While Search Console registrants do react faster, the effect is small relative to the 15-20% increase in remediation rate from Search Console messaging.

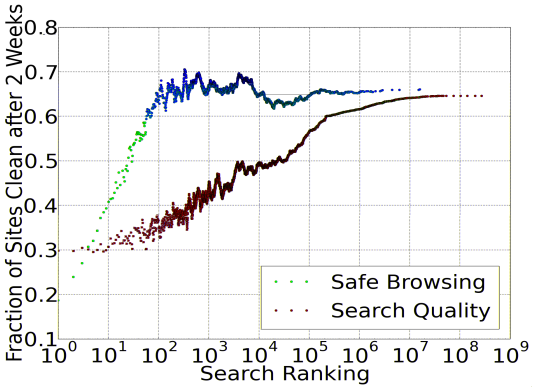


Figure 6: The percentages of sites clean after two weeks across different search rankings. Cleanup rate increases for higher ranked site, although the increase is stronger for Safe Browsing than Search Quality, possibly due to browser interstitials being stronger alert signals.

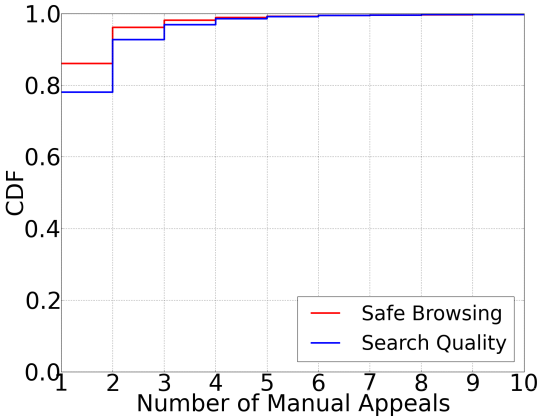


Figure 7: Number of manual appeals per hijacking incident. The majority of site operators successfully appeal on their first attempt.

5.7 Modeling Infection Duration

Given all of the potential factors that impact infection duration, we build a model to investigate which variables have the strongest correlation with faster recovery. We consider the following dimensions for each hijacking incident: detection source, Search Console usage, whether a Search Console message was sent, and all of a site's demographic data including age, size, ranking, and language. For site size and ranking, we use the log base 10. We exclude a site's software platform as we lack this annotation for over 86% of sites. Similarly, we exclude systemic versus localized infection labels as these annotations exist only for Search Quality. Finally, we train a ridge regression model [15] (with parameter $\lambda = 0.1$) using 10-fold cross validation on the first hijacking incident of each site and its corresponding remediation time, exclusively for sites that clean up. Ridge regression is a variant of linear regression that applies a penalty based on the sum of squared weights, where λ is the penalty factor. This brings the most important features into focus, reducing the weights of less important features.

Overall, our model exhibits low accuracy, with an average fit of $R^2 = 0.24$. This arises due to the limited dimensionality of our data, where sites with identical feature vectors exhibit significantly different infection durations. Despite the poor fit, we find Search Console alerts, search ranking, and Search Console registration exhibit the largest magnitude weights, with weights of -10.3, -6.1, and -3.3 respectively. This suggests that receiving a Search Console alert reduces infection lengths by over 10 days on average, a stronger effect than from other factors. Interpreting these results, we argue that direct communication with webmasters is the best path to expedited recovery, while popular websites naturally benefit from increased warning visibility and potentially more technically capable webmasters. Conversely, we find other factors such as the corresponding notification's language or a site's age and number of pages do not correlate with faster recovery. We caution these are only weak correlations for a model with high error rate, but as our prior analysis shows, they have the highest discriminating power at determining the lifetime of an infection.

6. WEBMASTER COMPREHENSION

Webmasters that receive a warning or notification must be technically savvy enough to contend with the corresponding security breach, or reach out to a third party that can help. Our dataset provides a lens into three aspects of webmaster comprehension: webmasters incorrectly requesting the removal of hijacking flags for their site when symptoms persist; sites repeatedly falling victim to new hijacking incidents in a short time window; and whether the duration that a site remains compromised improves after repeated incidents, a sign of learning over time.

6.1 Cleanup Attempts Before Successful

Both Search Console and Safe Browsing provide webmasters with a mechanism to manually appeal hijacking flags if webmasters believe their site is cleaned up. This triggers a re-scan or manual review by a Google analyst, after which the respective pipeline confirms a site is symptom-free or rejects the appeal due to an ongoing incident. We focus on webmaster-initiated appeals, as opposed to automated appeals that happen periodically, because the timestamp of a manual appeal denotes when a webmaster is aware their site is compromised and is taking action. Overall, 30.7% of Safe Browsing and 11.3% of Search Quality webmasters ever submit a manual appeal, of which 98.7% and 91.4% were eventually successful.

Figure 7 captures the number of webmaster cleanup attempts per hijacking incident before a site was verified symptom-free. We find 86% of Safe Browsing sites and 78% of Search Quality sites successfully clean up on their first attempt, while 92% of all site operators succeed in cleaning up within two attempts. Our findings illustrate that webmasters in fact possess the technical capabilities and resources necessary to address web compromise as well as to correctly understand when sites are cleaned. However, a small portion of the webmasters struggle to efficiently deal with compromise. For both Safe Browsing and Search Quality, at least 1% of the webmasters required 5 or more appeals.

For webmasters that fail at least one appeal, we measure the total time spent in the appeals process in Figure 8. We find the median Safe Browsing site spends 22 hours cleaning up, compared to 5.6 days for Search Quality. The discrepancy between the two systems is partially due to Search Quality requiring human review and approval for each appeal, as opposed to Safe Browsing's automated re-scan pipeline. Another factor is the fact that webmasters addressing Search Quality incidents tend to require more appeals than Safe Browsing. Our findings illustrate a small fraction of webmasters require a lengthy amount of time to manage their way through the appeal process, with some still struggling after 2 months. In the majority of these cases, a long period of time elapses between subsequent appeals, suggesting that these users do not understand how to proceed after a failed appeal or they give up temporarily.

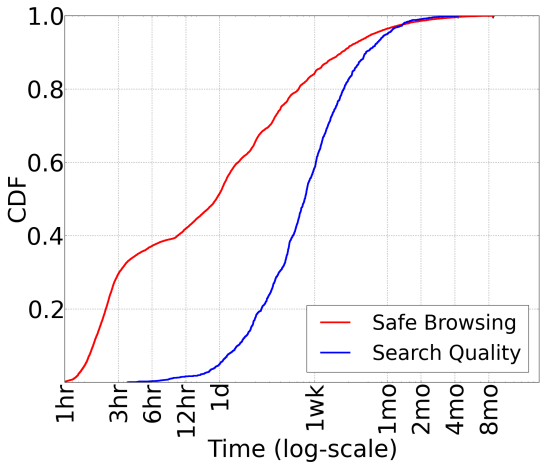


Figure 8: Total time webmasters spend cleaning up, exclusively for hijacking incidents where site operators submit multiple appeals.

6.2 Reinfection Rates

After webmasters remedy an infection, an important consideration is whether the operator merely removed all visible symptoms of compromise or correctly addressed the root cause. Webmasters that fail to remove backdoors, reset passwords, or update their software are likely to fall victim to subsequent attacks. Along these lines, we measure the likelihood that a previously compromised site falls victim to a second, distinct hijacking incident. To capture this behavior, we analyze the subset of all hijacking incidents that webmasters successfully cleaned and calculate what fraction Google flagged as hijacked again within one month. Given Safe Browsing and Search Quality detection operates on a per-day granularity, our measurement requires a minimum of a one-day gap between infection incidents.

We find that 22.3% of Search Quality sites and 6.0% of Safe Browsing sites become reinfected within 30 days after their first compromise incident. Figure 9 provides a CDF of the time gap between infections, restricted to those sites that become reinfected. We observe attackers reinfect over 10% of Safe BrowsingSasites and over 20% of Search Quality sites within a single day. After 15 days, this increases to 77% of Safe Browsing sites and 85% of Search

through which security services can communicate with webmasters, spurring remediation. Indeed, a majority of webmasters expressed that they were unaware of a hijacking incident until they were notified or saw a warning. Some webmasters requested that any site-level hijacking flag not take affect until one week after notification. However, such an approach both requires a direct notification channel (thus ruling out interstitials or search warnings) and also puts visitors at risk in the interim. These anecdotes highlight the duality of security when it comes to web compromise and the decision that web services must make in whose interests to prioritize.

7.2 Improving Remediation

Our study found that webmasters can significantly benefit from early notification of infections, but that webmasters fail to redress 40.5% of hijacking incidents. We offer three approaches for improving overall cleanup rates: increasing webmaster coverage, providing precise infection details, and equipping site operators with recovery tools.

Notification Coverage: Our findings showed that direct Search Console notifications outperformed global warnings, but that only 22% of Search Quality incidents and 32% of Safe Browsing incidents had a corresponding webmaster registered with Search Console. This stems in part from the requirement that webmasters must both know about Search Console and proactively register an account. While email may seem like an attractive alternative channel, we argue that identifying a point of contact remains the most significant hurdle. Previous studies relied on WHOIS abuse contacts, though it is unclear what fraction of recipients received or opened the notification [5, 6, 24]. For our study, we were unable to independently assess the efficacy of Safe Browsing's email notifications due to their tight coupling with browser interstitials. However, given that webmasters cleaned up 51% more Safe Browsing incidents when a Search Console email was triggered versus not, it is clear that WHOIS-based emails often fall into the void. While hosting providers are also well-positioned, the lack of a uniform protocol or API to alert hosted sites remains a barrier to adoption. Instead, we argue that notifications should expand to services with broader reach among webmasters such as social networks and analytics or ads platforms.

Report Content: A common theme among webmaster appeals was the desire for more detailed report logs of precisely what pages served harmful content. As our analysis of appeals found, 14–22% of webmasters lacked sufficient expertise or example code to clean up on their first attempt. As Vasek et al. previously showed, including more detailed reports expedites remediation [24]. Potential improvements might include screenshots of rogue pages, a tool for accessing a crawler's perspective of injected content, or more detailed diagnostic information. Alternatively, in the absence of a direct notification, browser interstitials and search warnings could include information targeted directly at webmasters rather than merely visitors. This is a tacit recognition that global warnings play a key role in recovery. That said, detailed logs may steer webmasters towards addressing symptoms of compromise rather than the root cause, yielding an increase in repeated hijacking incidents. We leave the investigation of how best to improve reports for future work.

Recovery Tools: While our findings demonstrate that many webmasters successfully recover, we are aware of few tools that help with the process. Such tools would decrease the duration of compromise and likelihood of reinfection. However, a question remains whether such tools generalize across attacks or not. Soska et al. found that attackers commonly exploit the same vulnerable software [19]. As such, it may be better to proactively notify sites of outdated software rather than wait till a hijacking incident, sidestepping the challenge of cleaning up specialized payloads.

7.3 Responsible Parties

The final challenge we consider is who should assume the responsibility for driving remediation. Site operators are best positioned to redress hijacking incidents, but our and prior work has shown that webmasters are often unaware their site is compromised until an outside alert. Alternatively, hosting providers own the serving infrastructure for compromised sites, of which security scanning could be a service. However, doing so comes at a financial cost or technical burden; today, few providers scan for harmful content or vulnerabilities [4]. Finally, ISPs, browser vendors, and search engine providers can enact incentives to spur action, but ultimately they possess limited capacity to directly help with remediation. These factors—representing the decentralized ideals of the Internet—make web compromise more challenging to address than account or credit card breaches where a centralized operator responds. The security community must determine which parties should bear the most responsibility for attending to compromised sites; what actions the parties should and should not pursue; and which mechanisms to employ to hold each accountable. Until then, compromise will remain an ongoing problem.

8. CONCLUSION

In this work, we explored the influence of various notification techniques on remediation likelihood and time to cleanup. Our results indicate that browser interstitials, search warnings, and direct communication with webmasters all play a crucial role in alerting webmasters to compromise and spurring action. Based on a sample of 760,935 hijacking incidents from July, 2014–June, 2015, we found that 59.5% of notified webmasters successfully recovered. Breaking down this aggregate behavior, we found Safe Browsing interstitials, paired with search warnings and WHOIS emails, resulted in 54.6% of sites cleaning up, compared to 43.4% of sites flagged with a search warning alone. Above all, direct contact with webmasters increased the likelihood of remediation to over 75%. However, this process was confounded in part by 20% of webmasters incorrectly handling remediation, requiring multiple back-and-forth engagement with Safe Browsing and Search Quality to reestablish a clean site. Equally problematic, a sizeable fraction of site owners failed to address the root cause of compromise, with over 12% of sites falling victim to a new attack within 30 days. To improve this process moving forward, we highlighted three paths: increasing the webmaster coverage of notifications, providing precise infection details, and equipping site operators with recovery tools or alerting webmasters to potential threats (e.g., outdated software) before they escalate to security breaches. These approaches address a deficit of security expertise among site operators and hosting providers. By empowering small website operators—the largest victims of hijacking today—with better security tools and practices, we can prevent miscreants from siphoning traffic and resources that fuel even larger Internet threats.

9. ACKNOWLEDGEMENTS

We thank the Google Trust and Safety teams—Safe Browsing and Search Quality—for insightful discussions and support in developing our data analysis pipeline. This work was supported in part by NSF grants CNS-1237265 and CNS-1518921. Opinions and findings are those of the authors.

10. REFERENCES

[1] Webmaster Tools now in 26 languages. http://googlewebmastercentral.blogspot.com/2008/05/webmaster-tools-now-in-26-languages.html, 2008.

[2] D. Akhawe and A. P. Felt. Alice in warningland: A large-scale field study of browser security warning effectiveness. In *Proceedings of the USENIX Security Symposium*, 2013.

[3] K. Borgolte, C. Kruegel, and G. Vigna. Meerkat: Detecting website defacements through image-based object recognition. In *Proceedings of the USENIX Security Symposium*, 2015.

[4] D. Canali, D. Balzarotti, and A. Francillon. The role of web hosting providers in detecting compromised websites. In *Proceedings of the 22nd International Conference on World Wide Web*, 2013.

[5] O. Cetin, M. H. Jhaveri, C. Ganan, M. Eeten, and T. Moore. Understanding the Role of Sender Reputation in Abuse Reporting and Cleanup. In *Workshop on the Economics of Information Security (WEIS)*, 2015.

[6] Z. Durumeric, F. Li, J. Kasten, N. Weaver, J. Amann, J. Beekman, M. Payer, D. Adrian, V. Paxson, M. Bailey, and J. A. Halderman. The Matter of Heartbleed. In *ACM Internet Measurement Conference (IMC)*, 2014.

[7] M. Egele, G. Stringhini, C. Kruegel, and G. Vigna. COMPA: Detecting Compromised Accounts on Social Networks. In *Proceedings of the Network and Distributed System Security Symposium (NDSS)*, 2013.

[8] A. P. Felt, A. Ainslie, R. W. Reeder, S. Consolvo, S. Thyagaraja, A. Bettes, H. Harris, and J. Grimes. Improving SSL Warnings: Comprehension and Adherence. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015.

[9] Google. Fighting Spam. http://www.google.com/insidesearch/howsearchworks/fighting-spam.html, 2015.

[10] Google. Googlebot. https://support.google.com/webmasters/answer/182072?hl=en, 2015.

[11] Google. Safe browsing transparency report. https://www.google.com/transparencyreport/safebrowsing/, 2015.

[12] Google. Search Console. https://www.google.com/webmasters/tools/home?hl=en, 2015.

[13] Google. "This site may be hacked" message. https://support.google.com/websearch/answer/190597?hl=en, 2015.

[14] Google. "This site may harm your computer" notification. https://support.google.com/websearch/answer/45449?hl=en, 2015.

[15] A. E. Hoerl and R. W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 1970.

[16] M. Jones. Link Shim - Protecting the People who Use Facebook from Malicious URLs. https://www.facebook.com/notes/facebook-security/link-shim-protecting-the-people-who-use-facebook-from-malicious-urls, 2012.

[17] M. Kührer, T. Hupperich, C. Rossow, and T. Holz. Exit from hell? reducing the impact of amplification ddos attacks. In *Proceedings of the USENIX Security Symposium*, 2014.

[18] N. Provos, P. Mavrommatis, M. A. Rajab, and F. Monrose. All your iFRAMEs point to us. In *Proceedings of the 17th Usenix Security Symposium*, pages 1–15, July 2008.

[19] K. Soska and N. Christin. Automatically detecting vulnerable websites before they turn malicious. In *Proceedings of the USENIX Security Symposium*, 2014.

[20] StopBadware. Request A Review. https://www.stopbadware.org/request-review, 2015.

[21] StopBadware and CommTouch. Compromised Websites: An Owner's Perspective. https://www.stopbadware.org/files/compromised-websites-an-owners-perspective.pdf, 2012.

[22] K. Thomas, F. Li, C. Grier, and V. Paxson. Consequences of connectivity: Characterizing account hijacking on twitter. In *Proceedings of the Conference on Computer and Communications Security*, 2014.

[23] Twitter. Unsafe links on Twitter. https://support.twitter.com/articles/90491, 2015.

[24] M. Vasek and T. Moore. Do Malware Reports Expedite Cleanup? An Experimental Study. In *USENIX Workshop on Cyber Security Experimentation and Test (CSET)*, 2012.

[25] M. Vasek and T. Moore. Identifying risk factors for webserver compromise. In *Financial Cryptography and Data Security*, 2014.

[26] D. Y. Wang, S. Savage, and G. M. Voelker. Cloak and dagger: dynamics of web search cloaking. In *Proceedings of the ACM Conference on Computer and Communications Security*, 2011.