Vision: What the hack is going on? A first look at how website owners became aware that their website was hacked

Anne Hennig
Karlsruhe Institute of Technology
Kalrsruhe, Germany
anne.hennig@kit.edu

Nhu Thi Thanh Vuong
Karlsruhe Institute of Technology
Kalrsruhe, Germany
nhu.vuong@student.kit.edu

Peter Mayer
South Denmark University
Odense, Denmark
mayer@imada.sdu.dk

ABSTRACT
Websites are an essential part of today’s business activities. Content Management Systems (CMS) are known for the fact that even laypersons can create good-looking websites with simple means and without huge costs. But if websites are not maintained regularly, they are prone to vulnerabilities. Such vulnerabilities can be abused, e.g., for third party redirects. Informing website owner about this type of attack is challenging. To gain more information about how website owners are informed about vulnerabilities on their websites, we invited 156 website owners to participate in an online survey. We asked those who had fixed the third party redirect before we could inform them, how they became aware of the attack. The participants could choose to answer the questionnaire via a link to an online platform, or to send their answers back to us via e-mail. Only 11 people answered our questionnaire, and only four people were already aware of the attack before our invitation e-mail. Based on these four answers, we assumed that we can confirm previous research with respect to the design of a vulnerability notification. Nevertheless, it would be interesting to see if – with a bigger sample – we can also confirm our findings that a) online surveys, even if they can only be accessed by clicking an unknown link, are preferred over responding via e-mail, b) the number of responses can be increased by sending out several reminder, and c) a sender attributed with higher authority increases the response rate. Furthermore, we suggest that future research on vulnerability notifications questions the use of the term trustworthiness, and examines whether recipients distinguish between credibility and trustworthiness of notifications when remediate
ing attacks.

CCS CONCEPTS
• Security and privacy → Social aspects of security and privacy

KEYWORDS
vulnerability notification, website hacking, online survey, credibility vs trustworthiness

1 INTRODUCTION
Nowadays websites are a “must-have” for business owners. Not only is e-commerce becoming a vital part of business activities [16]. Also providing information like opening hours or contact information for customers is essential for being visible in the first place. In 2023, more than two third of websites were based on Content Management Systems (CMS) [21]. While those provide good-looking and easy-to-use website solutions for laypersons [15], the effort for maintenance is often neglected. According to recent data by Kasturi et al. [8], over 47,000 plugins for popular CMS which were installed on nearly 25,000 websites contained malicious code, making websites vulnerable for a number of different attacks [2]. The 2022 Website Threat Research Report [7] showed that vulnerable plugins are a common attack vector: Around one-third of hacked websites in their sample contained at least one vulnerable plugin or theme.

While most attacks are immediately noticeable by website owners (e.g. defacement, or spread of malware), some rather unknown attacks are not easy to recognize. Within a current research project, we found websites that redirect to malicious websites, e.g. fake online pharmacies, fake online warehouses, fake online casinos, or porn websites. To install the redirects in the code of the website, an attacker needs access to the webspace (e.g. through leaked passwords or by exploiting vulnerabilities). The manipulation is not visible on the genuine website, but the search engine results list malicious URLs for this website. By clicking on such a malicious URL that is added to the code of the genuine website, the user is redirected to the fake website. While this might be annoying for the user, it can cause reputational damage to the website owner. More important, third-party redirects are a sign that an attacker gained access to the webspace. Which can then result in further attacks if the redirect is no longer profitable (e.g. defacement, or spread of malware).

The malicious URLs can be identified by crawling search engine entries. For our research, we use AI to identify URLs that do not match the corresponding domain and show signs of third-party redirects. The overall goal is to effectively inform website owners about the manipulation, and make them aware of the severity of this kind of attack. But we learned that for some websites we could not identify a third-party redirect anymore before sending out notifications. We assume that those websites have already remediated

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

EuroUSEC 2023, October 16–17, 2023, Copenhagen, Denmark

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-0814-5/23/10...
$15.00
https://doi.org/10.1145/3617072.3617101
the third-party redirect, but we have no further information on how the website owners were informed about the attack, or if they were informed at all. Thus, the goal of this research is to answer the following research questions:

**RQ1** How did website owners become aware of the third-party redirect or who notified them about the attack?

**RQ2** How trustworthy\(^1\) did the website owners perceive the respective notification?

**RQ3** Which notification channels would website owners deem trustworthy in case of future notifications?

By answering our research questions, we will provide comprehensive insight into (a) how website owners are notified about vulnerabilities and (b) what their reasoning is to perceive certain notification channels as trustworthy. Previous research already reports on the trustworthiness of notifications, but to the best of our knowledge, no study has yet surveyed victims that had already remediated a vulnerability before being notified.

In previous studies, website owners should report on their experiences and assess their perception of vulnerability notifications retrospectively and based on a notification that has previously been sent \([3, 6, 10, 13, 18, 20, 23, 24]\). We assume that previous surveys are thus biased: When asked to judge, for example, the trustworthiness of a notification, they will be influenced by the notification they have seen before (Priming) and assess the trustworthiness accordingly. We also wanted to get an estimate of how many website owners detected a vulnerability themselves, how many were notified by a third party (e.g. a customer or an employee), or if there are other factors that lead to the remediation (e.g. website has been taken down in the meanwhile without the vulnerability being detected by anyone). Therefore, we deem it necessary to ask for other channels of notification – including detection by the website owners themselves – and how these notifications were perceived, without the context of a previously sent notification.

We report on the related work in section 1, and present the methodology of our survey design in section 2. We report the results of our study in section 3, and discuss our challenges to reach out to website owners and provide recommendations for future work in section 4.4. We then conclude our paper with a summary in section 5.2.

## 2 RELATED WORK

We already know from previous research how notifications are perceived that have been sent out by the researchers to notify website owners about a certain vulnerability or misconfiguration. These studies were mainly designed as experimental studies, measuring the ratio of remediated websites \([3, 9–13, 18–20, 22–27]\). Some of these experiments were accompanied by quantitative surveys \([3, 10, 13, 18, 20, 23, 24]\) or qualitative interviews \([6]\) to find out how notifications were perceived by the website owners.

The main questions in these studies were whether website owners were aware of the vulnerability prior to the notification; how website owners perceived the respective notification with respect to e.g. trustworthiness; how the respective notifications could be improved; why the problem has not been remediated; and whether the website owners want to receive further notifications alike.

To the best of our knowledge, our study is the first to survey website owners who seem to have remediated the attack prior to a notification. Thus, our study extends previous findings on vulnerability notifications.

## 3 METHODOLOGY

Based on surveys from previous notification studies \([10, 13, 23, 24]\), we designed an online survey to answer our research questions. To identify relevant participants we compiled a list of websites from Germany and Austria that were found affected by a third-party redirect between September 2020 and February 2022, and that did not show redirects to malicious websites in July 2022 anymore. Since the only relevant information for us was whether a website was affected by the third-party redirect, we did not further investigate the companies behind the websites, and can, thus, not make any statements about the scope of our sample. From the broader project context, we know that mainly – but not exclusively – SMEs as well as associations and freelancers are affected by these attacks.

As a first step, the contact information for the websites were manually\(^2\) extracted from the imprint or – in case websites did not have an imprint – the contact page of the website. We were especially looking for the name and the e-mail address of the website owners to make the invitations as personal as possible, as suggested by Hennig et al. \([6]\). We excluded websites where we could not find any contact information on the website.

We choose invitations via e-mail since this has been evaluated as the most cost-efficient \([13, 18]\) and widely accepted \([6]\) way to contact website owners, although we acknowledge that letters \([12, 17]\) or phone calls \([5, 6]\) might result in higher response rates. The design of the e-mail invitation was based on the recommendations given by Maass et al. \([14]\) and Hennig et al. \([6]\).

To lower the threshold for answering our questions, we also included the questions from the online survey within our invitation e-mail. Including a questionnaire in the e-mail was also applied by Hennig et al. \([5]\), since previous studies found that recipients are suspicious of unknown links in e-mails (in the context of vulnerability notifications e.g. \([6, 13, 19]\)). Furthermore, including the survey questions in the invitation e-mail gives participants the opportunity to elaborate on their answers and provide more details \([5]\). Instead, in the online survey we mainly used closed questions to simplify answering the questions as much as possible. Also, the answers in the online survey were collected anonymously, i.e., we did not include personalized links or collect personal data (e.g. names, domain names, or similar).

By offering both – answering the questions via the online survey, and as a reply to our invitation e-mail – we aimed at attracting more participants: Those who appreciate the anonymity of an online survey and trust the link, and participants who appreciate a lower effort option to answer our questions but might mistrust the link.

\(^{1}\)Note, that we use the term ‘trustworthy’ here in accordance with previous research. We elaborate in the discussion section why we think this should be questioned in following research.

\(^{2}\)Previous research could show that automatic extraction from WHOIS data or creating generic e-mail addresses causes high bounce rates \([13, 18, 19, 24]\). Thus, we followed the recommendations of \([6]\) and \([5]\) and extracted contact data manually.
We asked the website owners to answer the following questions:

Q1 Were you at any time aware of the third-party redirect to a fake shop on your website? [yes / no]
Q2 How did you become aware of this third-party redirect? For example, did you discover it yourself, was it discovered by an employee, or were you perhaps informed by a third party?
   (a) If you were informed by a third party, who exactly brought it to your attention? Was it, for example, your external service provider or your hosting provider?
   (b) If you were informed by a third party, how were you informed about the third-party redirect? Was it by e-mail, phone call, or personal approach?
   (c) If you were informed by a third party,
       • Which aspects of the notification did you find most trustworthy?
       • Which motivated you to take the problem seriously and fix it?
       • Which aspects of the notification were NOT trustworthy for you?
       • Which exactly would you like to see in future vulnerability notifications?
Q3 Which information was especially helpful for you to identify and remediate the problem?
Q4 Would you like to keep notified by qualified third parties about vulnerabilities on your website? [yes / no]
   (a) If yes, how would you like to be informed about vulnerabilities on your website in the future? [e-mail / letter / phone call / other]

We pre-tested our study design with 18 participants recruited from friends or relatives of one of the authors knowingly; that feedback from friends and relatives might be more favorable and biased. Our goal was to mainly test the technical functionality and the comprehensibility of the instructions. Thus, "wrong" answers would have been an indicator that the questions were not comprehensible. Participants were asked to evaluate the wording of the invitation e-mail, and go through both questionnaires (online and reply to our invitation e-mail). All participants considered the wording of the e-mail invitation, and the questionnaires to be good. No spelling errors or technical malfunctions were detected. Since we received satisfactory answers in each case, we assume that the procedure and the questions in the questionnaire were understood.

In the end, 156 invitation e-mails were sent from an e-mail address related to the research project on August 2, 2022. The website owners were asked to answer our questions by August 30, 2022. After three weeks, on August 23, 2022, an e-mail reminder was sent. On August 30, 2022, a second reminder was sent with an extension of the deadline until September 13, 2022. Furthermore, a preface was added and this reminder was sent from a different e-mail address, one aligned with the authors’ institution and not the research project. A third reminder e-mail from this sender was sent on September 13, 2022, and the deadline was extended for a second time until September 20, 2022.

4 RESULTS

4.1 Participant Sample
Out of 156 invitation e-mails that were sent out on August 2, 2022, nine e-mails could not be delivered. Thus, the total number of recipients was reduced to 147. Of these, a total of eleven website owners answered our questions.

Eight website owners responded as a reply to our invitation e-mail: One rejected to participate in the survey, four answered the questions in the e-mail, and three responded with questions themselves, but did not answer our questions via e-mail. It remains unknown whether these participants may have completed the online questionnaire. The links to the online survey were not personalized, thus, we cannot be entirely sure if one of those who came back with questions answered our online survey in the end.

The online survey was accessed 26 times in total. Eight participants started the online questionnaire. Of these, one person refused to give consent. Seven participants answered the online questionnaire, which adds up to eleven completed questionnaires in the end. Again, we cannot be entirely sure that people who answered our questions via e-mail also answered the questions in the online survey and vice versa. However, after checking the data sets, we found the answers to be different enough to inspire confidence that all responses were from different participants, and we assume that we obtained eleven unique answers.

We obtained two completed questionnaires after the initial e-mail invitation (one answered via e-mail, one answered the online questionnaire), and two more after the first reminder (both answered the online questionnaire). After the second reminder we obtained four completed questionnaires (two answered the questions via e-mail, and two the online questionnaire), and three more after the third reminder (one answered the questions via e-mail, two answered the online questionnaire). Due to the low number of completed questionnaires, we only descriptively report on the answers in the following. We report count data instead of percentages to avoid over-generalizing.

4.2 RQ1: How did website owners become aware of the third-party redirect or who notified them about the attack?

When asked whether they were aware of the attack (Q1), seven participants indicated that they were not aware of the third-party redirect prior to our invitation e-mail. The remaining four participants were aware of the third-party redirect. All four described different ways of notification (Q2 a & b): One was informed by an employee who recognized problems with the website; one was informed by a third party, which was described as the hosting provider, serving as the external service provider for the company’s website; one fell victim of a ransomware attack, and became aware of the redirect after the ransom e-mail. This participant further recognized that spam e-mails were sent out via their e-mail account. One recognized the attack themselves because they could not log in to the admin account of the website anymore. After our e-mail invitation, they found a notification by their hosting provider about a possible virus infection.

---

The questions were slightly different in the online survey since we could include filter questions and, therefore, lead the participants to the relevant questions according to their answers in previous questions.
4.3 RQ2: How trustworthy did the website owners perceive the respective notification?

When asked to name aspects that were found trustworthy, resp. not trustworthy (Q2 c), two persons referred to the sender of the notification. One person said a hosting provider can probably be considered trustworthy. The second person said, for them, information from business partners, patients, or the hosting provider would be trustworthy, whereas all e-mail addresses that sound strange are not trustworthy. They also added that they considered the initial notification to be phishing since they were asked to click on a link. The third participant did not assess the trustworthiness themselves. They had no knowledge about the topic, so they simply forwarded the notification to an external service provider. Since this website owner had been informed by the attackers, they further said, the fact that an attack has taken place was absolutely credible, but the hackers as informants cannot logically claim any trustworthiness.

We also asked for aspects in the notification that motivated the participants to take the problem seriously and fix it (Q2 c) and which information were helpful to remediate the attack (Q3). One person answered that they did not want to administer a hacked website because they do not want to harm others. Another responded that data safety is important for them, even if they think in this case, the possibility of data theft was relatively small. The remaining two participants said they still haven’t understood the problem. One said they rebuilt the site based on a backup, but have no information on what had happened. The other said they were struggling to remediate the third-party redirect, and admitted that they still have no clue what had happened.

4.4 RQ3: Which notification channels would website owners deem trustworthy in case of future notifications?

When asked whether the participants would like to be informed about vulnerabilities in the future (Q4), eight participants agreed and three denied. We also asked which notification channels the participants prefer. Five participants answered they would prefer e-mail notifications. One answered they prefer a combination of e-mail and letter, another one said they prefer a combination of e-mail and phone call. One person preferred the customer account of the hosting provider.

When asked for the content of future notifications (Q2 c), one person answered that the sender should be reasonable and trustworthy, and the notification should contain the following: a) information about what exactly had happened, b) since when the redirect is likely to be active, c) what the potential negative side effects of the attack are – for the website owner as well as for the general public, and d) if possible provide best practices that help remediate and prevent such attacks. Another person said that they would like to have a notification from the hosting provider, and also more support from the hosting provider to understand the scope of the problem and to get a permanent solution.

5 DISCUSSION

The goal of our research was to gain insight into how website owners became aware of vulnerabilities on their website. Furthermore, we wanted to extend the body of literature on the perception of trustworthy vulnerability notifications.

Unfortunately, the small number of completed questionnaires makes it impossible to draw generally valid conclusions. We reached a response rate of 7.48%, which is comparable or even better to previous research on that topic (e.g. [10, 13, 18, 23, 24]). But due to the small sample, we could only analyze eleven completed questionnaires. Nevertheless, our results provide relevant insights.

5.1 Discussion of the Results

RQ1: How did website owners become aware of the third-party redirect or who notified them about the attack? We asked website owners if they were aware of the third party redirect on their website, and if yes, how they first became aware. Only four participants were aware of the redirect prior to our invitation e-mail. Two were informed by their hosting provider, one was “informed” by the attacker who blackmailed them, and one was informed by an employee. Due to the small number of responses, we cannot draw a conclusive picture here. However, from these responses, we can see that a) some hosting providers seem to be aware of the attack and are able to recognize it, and b) in some cases the attackers are not satisfied with just a redirect but change their tactics and implement a ransomware attack, for example.

Interestingly, the majority of the respondents were not aware of the third-party redirect prior to our invitation e-mail. This raises the question, why these websites were found remediated. We assume that either our detection tool found false positives, or the redirect was removed externally. Maybe the malicious website was closed by the attacker because it was not profitable, or due to criminal prosecution. If the redirect was removed by the attacker, the vulnerability itself still exists and the website could be exploited for further attacks. Unfortunately, we cannot verify whether this is the case.

RQ2: How trustworthy did the website owners perceive the respective notification? To answer our second research question, we asked the participants who received a notification, which aspects of the previous notification they found most trustworthy, which aspects were not trustworthy for them, what motivated the website owners to take the problem seriously and remediate it, what exactly they would like to see in future vulnerability notifications, and which information was especially helpful for them to identify and remediate the problem.

Again, due to the small number of responses (only three participants answered these questions), we cannot make general statements. It seems that the sender of a vulnerability notification has an impact on the trustworthiness, as shown in previous research (e.g. [6, 13, 18, 24]). Two participants named the hosting provider as a trustworthy sender, and one also added business partners or patients as trustworthy senders - all senders with whom a (business) relationship exists. As proposed in Hennig et al. [6], an existing relationship with the sender seems to be generally important. Thus, the decision of whom website owners trust depends a lot on the individual. One participant explicitly mentioned that clicking on a link was not trust-promoting for them. Here, we can confirm previous recommendations (e.g. by [6, 14]) to not include links in vulnerability notifications.
We also asked what aspects in the notification motivated the participants to take the problem seriously. While two participants said they didn’t want their website to be a threat to others, two participants admitted that they still have not understood the problem. We assume, that if the problem is understood, the main worry is that others might be harmed. Thus, adding consequences of an attack in a vulnerability notification may increase motivation.

Interestingly, one of our participants distinguished between trustworthiness and credibility: The notification by the hacker was credible, although the sender of the notification was not trustworthy at all. This reflects that trustworthiness is only a part of a successful notification. Fogg & Tseng [4] summarized that persons assess credibility based on perceived trustworthiness and perceived expertise. In the context of vulnerability notifications this would mean that rather than trustworthiness the perceived credibility of a notification results in remediation. Thus, a notification content or a sender that conveys a high level of expertise could outweigh trustworthiness, for example. Fogg & Tseng [4] propose further factors that affect computer credibility as well as different models of computer credibility evaluation. To the best of our knowledge, none of the related studies on vulnerability notifications have discussed the distinction between trustworthiness and credibility. Therefore, it would be highly interesting to further investigate this relationship specifically in the context of vulnerability notifications.

RQ3: Which notification channels would website owners deem trustworthy in case of future notifications? To answer our third research question, we asked participants who were informed about the third-party redirect what they would like to see in future vulnerability notifications. We also asked all participants if they would like to be notified by qualified and reputable third parties about vulnerabilities on their website in the future, and if so, which channels they prefer.

The majority of our participants agreed that they want to be informed about vulnerabilities on their website. Also, most participants preferred to be notified via e-mail, or a combination of letter / phone call and e-mail. By tendency, these results seem to be in line with previous research [3, 6, 10, 13, 18, 24].

With respect to recommendations for the content of future vulnerability notifications, we can only take two meaningful answers into consideration. One participant asked for detailed information about the redirect and potential side effects, how to remediate it, and how to protect from future attacks. Another participant also mentioned that they would like to have more support. They held the hosting provider responsible for providing support opportunities.

Previous research already suggested providing incentives for remediation [6, 13, 14, 22, 24], providing a clear description of the attack and point out consequences [6], and raising awareness among hosting providers [1, 25]. Both of the responses we got confirm these recommendations.

5.2 Discussion of the methodology

We contacted 156 website owners via e-mail, and we received eleven completed questionnaires after three reminder e-mails. Seven participants responded to the online survey and four answered our questions anonymously via the link to the online survey. So it might be that while links in e-mails are deemed suspicious in general, links to surveys are accepted as soon as the notification is deemed credible. Another explanation would be that people are increasingly used to answer online surveys.

It would also be interesting to investigate the effect of reminders in more detail. We received two completed questionnaires each after the initial invitation e-mail and the first reminder. We received four completed questionnaires after the second reminder, and three after the third reminder. It seems to be promising to send out several reminders since the number of responses continuously increased after each e-mail. Since the second and third reminders were sent by another e-mail address, the increase in responses could also indicate that more weight was given to a more senior sender with a clear affiliation to a research institution. To prove statistical significance, these findings need to be confirmed with a bigger sample.

6 CONCLUSION

Although well researched, the field of vulnerability notifications has not yet found the best way to effectively notify website owners about vulnerabilities on their websites. We already know from previous notification experiments how website owners perceive a previously sent notification with respect to, e.g., trustworthiness, how notifications could be improved, why issues have not been remediated after a notification, and whether the website owner wants to receive future notifications. To the best of our knowledge, all results are based on notifications that have previously been sent out to the website owner. We argue, that the website owners might be influenced by a certain type of notification (priming effect), and the results might, therefore, be biased.

Thus, we wanted to survey website owners who were probably notified about an attack by an unknown third party, and not as part of a notification experiment. To answer our research questions we designed an online survey and invited 156 website owners to participate in our survey. We offered them two options to answer our questions: They could either click on a link and answer the questions anonymously via the link to the online survey. By offering these two options we hoped to address participants who appreciate anonymity but trust a link to an online survey, and participants who appreciate a low-threshold option to answer our questions via e-mail but mistrust a link.

In total, 16 website owners responded to our invitation e-mail and we received eleven completed questionnaires: seven participants responded to the online survey and four answered our questions as a reply to our e-mail. We sent three reminder e-mails using two different e-mail addresses. Unfortunately, the number of responses was too small to analyze the results in detail. But it would be interesting to see if, with a bigger sample, we can confirm the findings that a) online surveys – even if they can only be accessed by clicking an unknown link – are preferred over responding via e-mail, b) the number of responses can be increased by sending out several reminders, and c) a sender attributed with higher authority increases the response rate.
Besides choosing a bigger sample, it would be interesting to analyze the remediated websites in more detail. Seven out of eleven participants were not aware of the third-party redirect prior to our invitation e-mail. One explanation is that our detection tool found too many false positives. The alternative would be to manually review all findings, which causes huge effort depending on the sample size. Another explanation is that the attacker themselves removed the redirect or the target website was taken down. In either case, the vulnerability and the unauthorized access would still exist, and the website could be exploited for further attacks.

With respect to the design of a vulnerability notification, we can confirm previous research. Although the sender on a notification seems to play a major role for website owners, the decision of whom website owners trust varies, and it seems that there is not a trustworthy sender. Also, one of our participants raised an interesting issue: By stating that they remediated the problem not because they found the notification trustworthy, but because it was credible to them. "Credibility" was mentioned in previous vulnerability notification studies, but it was solely used synonymously to "trustworthiness", and it was not discussed whether participants distinguished between trustworthiness and credibility when assessing if a problem needs their attention [22, 26, 27]. We suggest to further investigate the differences between trustworthiness and credibility specifically in the context of vulnerability notifications. Maybe this helps answering the question, why there is not a trustworthy sender. Also, one of our participants asked for a clear description of the problem being reported. Another of our participants complained that they had not received effective help from their hosting provider to remediate the problem. Maybe this helps answering the question, why there is not effective help from their hosting provider to remediate the problem.

ACKNOWLEDGMENTS

This research is supported by the Federal Ministry of Education and Research, Germany, as part of the BMBF INSPECTION project (Zuwendungszummer 16KIS1113), and it is supported by funding from the topic Engineering Secure Systems, topic 46.23.01 Methods for Engineering Secure Systems, of the Helmholz Association (HGF) and by KASTEL Security Research Labs.

REFERENCES


317