POSTER: Mapping the Landscape of Large-Scale Vulnerability Notifications

Ben Stock
CISPA, Saarland University
Saarland Informatics Campus
stock@cs.uni-saarland.de

Giancarlo Pellegrino
CISPA, Saarland University
Saarland Informatics Campus
gpellegrino@cispa.saarland

Christian Rossow
CISPA, Saarland University
Saarland Informatics Campus
crossow@mmci.uni-saarland.de

Martin Johns
SAP SE
martin.johns@sap.com

Michael Backes
CISPA, Saarland University &
MPI-SWS
Saarland Informatics Campus
backes@mpi-sws.org

ABSTRACT

The Internet is an ever-growing ecosystem with diverse software and hardware applications deployed in numerous countries around the globe. This heterogenous structure, however, is reduced to a homogenous means of addressing servers, i.e., their IP address. Due to this, analyzing different Internet services for vulnerabilities at scale is easy, leading to many researcher focusing on large-scale detection of many types of flaws. On the other hand, the persons responsible for the administration of said services are as heterogenous as the Internet architecture itself: be it in spoken languages or knowledge of technical details of the services.

The notification of vulnerable services has long been treated as a side note in research. Recently, the community has focussed more not only the detection of flaws, but also on the notification of affected parties. These works, however, only analyze a small segment of the problem space. Hence, in this paper, we investigate the issues encountered by the previous works and provide a number of future directions for research, ultimately aiming to allow for an easier means of notifying affected parties about vulnerabilities at scale.

Introduction

The last years of research have produced a multitude of tools which allow us to analyze large numbers of services for different types of vulnerabilities. These range from infrastructure-level flaws such as Heartbleed [2] or NTP amplification vulnerabilities to application-level bugs such as Client-Side Cross-Site Scripting. While most of the research has primarily focussed on means of detecting such bugs at scale, little attention has been given to the process of effectively notifying affected parties. In the last year, this new area of the research space has been partially explored by both our own work [6] and the work of others [1, 4]. These works, however, only investigated a small fraction of the problem space. More precisely, they have shown that using the current infrastructure available to researchers, it is not feasible to conduct notifications at scale. Therefore, in this paper we analyze the problem space in more detail and highlight which technical and human factors come into play when trying to disseminate vulnerability information at scale.

Background on Vulnerability Disclosures

While the research community has become more proficient in discovering vulnerabilities at scale, the notification of affected parties has mostly been treated as a side note, e.g., by **(author?)** [3] and **(author?)** [2]. In addition, additional research has been focussing on the notification of *infected* rather than *vulnerable* sites [1, 5]. More recently, both our own work [6] as well as concurrent research by **(author?)** [4] have analyzed the process of notifying vulnerable services at scale while measuring the exact impact of different variables, e.g., communication channels or languages.

In our work, we disclosed over 35,000 vulnerable Web sites to different parties. Next to Trusted-Third Parties (i.e., CERTs and a trusted mailing list) and hosting providers of the vulnerable sites, we also included direct contacts which use the affected *domain* as an anchor. To that end, we notified the domain's registrant and send emails to generic aliases (e.g., info@ or security@) for the domain. For the lookup process of the domain registrant, we used information provided from WHOIS entries for each domain. This data, however, is inherently incomplete: for 18.5% of all domains in our data set, we could not find a point of contact. Even though our results were statistically significant, the fix rate was unsatisfactory: at the end of our experiment, almost 75% of notified domains were still vulnerable.

Our data set consisted of two types of flaws: well-known WordPress vulnerabilities and previously-unknown Client-Side XSS flaws. While the WordPress installations were contained in the Top 1M sites, the Client-Side XSS were restricted to the Top 10,000 sites. We discovered that the same notification channels showed interesting differences between the data sets: while CERTs performed best for the

CCS'16 October 24-28, 2016, Vienna, Austria

ACM ISBN 978-1-4503-4139-4/16/10.

DOI: http://dx.doi.org/10.1145/2976749.2989057

high-ranked sites, they performed comparatively bad for the average WordPress installation. This is likely due to the priority given by the CERTs to high-ranked Web sites over Top 1M sites. We found that the largest CERT in our data set, responsible for well over 50% of all domains, only reacted after the end of our campaign. Similarly, the other indirect channel, i.e., the providers, performed worst for high-ranked Web sites. This can in parts be explained by the fact that the top 5 providers (accounting for approximately 25% of all domains) did not react upon our notification at all, thereby stopping the notification effort dead in its tracks.

In concurrent work, **(author?)** [4] notified administrators about infrastructure-level vulnerabilities. To that end, they notified CERTs as well as hosting provider abuse contacts about publicly accessible industry control systems (ICS), improperly configured IPv6 firewalls for dual-stack hosts, and servers susceptible to be used in amplification attacks. In total, they notified approximately 6,500 entities of the vulnerabilities they had discovered. To look up the corresponding hosting providers, they also utilized the WHOIS protocol for the IPs of the vulnerable systems. Depending on the type of vulnerability, they found that up to 20% of the hosts in question did not have any information on the abuse contact for the specific IP range, which underlines the issues we faced for domain owners. Nevertheless, the hosting providers performed better than the CERTs.

For ICS and IPv6 they found that their notification campaign improved upon the fix ratio in a statistically significant manner. However, they found that on top of the control group, only 11% of the notified contacts fixed the flaws. Moreover, for the amplifiers, no significant improvement could be observed. Apart from these results, they also experimented with different message verbosity and a link to follow-up information. They found that messages describing the problem in detail in the initial email worked best. The authors also tried to notify parties in their native language. Contrary to the intuition, this led to a lower fix rate. The paper presents anecdotal evidence that a, e.g., German message from a US institution rather inspired distrust. For the specific channels they used, they discovered that contacting the hosting providers worked best, whereas their own CERT, i.e.. US CERT, did not act upon their notification at all.

These two works have taken a first look into the problem space of notification at scale. They come to the same, unsatisfactory conclusion: the impact on the general vulnerability population is very low. Moreover, no long-term benefits could be observed. In an analysis of the patch behavior of WordPress installations, we discovered that the average time to installing a security patch for sites which previously acted upon our notification was only slightly lower than in the control group. We argue that these works have provided valuable insights into the problems researchers are faced with. In the following, we therefore discuss open questions both on technical and human aspects resulting from the works and show that the community has only scratched the surface of the notification problem space.

Future Research Directions

In the following, we discuss a number of technical and human challenges which have to be addressed to allow for more successful notifications in the future.

Technical Challenges

Dedicated Security Contacts— The previous works have shown that a number of technical challenges have to be overcome to allow for successful notifications at scale. Both works in parts relied on WHOIS to determine contact points, but could not do so for up to 20% of affected parties. On top of that, the WHOIS information for IP ranges typically only features an *abuse* contact, not a *security* contact. Although depending on the provider, the abuse and security team might overlap, we found evidence that our *vulnerability* notification were misconstrued as abuse complaints, resulting in providers threatening their customers with account deletion. Hence, we argue that a dedicated security contact should be established for IP ranges, such that vulnerability notifications can reach the correct contact.

New Communication Channels— In our study, we directly notified domain owners and generic email aliases for the domain in question, i.e., we sent one (owners) and four (generic) emails for each domain, respectively. At the same time, we conducted a *large-scale* analysis, i.e., we notified almost 18,000 domains this way. The emails we received throughout our campaign indicate that this massive mail campaign sometimes caused issues with spam filtering. Even though we took the necessary precautions in setting up our mailserver, such issues cannot always be prevented. We therefore find that sending emails to notify large numbers of vulnerable parties is far from optimal and new means of communication channels must be researched. We envision that this could be done by centralized authorities. Such an authority could establish trust in a researcher once, allowing him to then use the infrastructure to notify site owners. An example of such an infrastructure is the Google Search Console, which allows domain owners to register their domains to subsequently get notifications on issues detected on their sites via the Web interface [5]. In this case, however, access to researchers is limited, as only Google can access this communication channel to reach out to affected parties.

Trustworthiness of Channel— In general, one issue which can severely impair the success of a notification campaign is the trust in the disclosed information. In email communication, such trust can be established in using signed emails, for which the recipient can verify the chain of trust from a root to the sender. In general, however, we cannot assume that all email clients correctly display signatures (e.g., Web mailers) and moreover that the recipient is aware of the concept of signed emails. In an experiment we conducted after the end of our study, we found that sending emails containing all details of the vulnerability lead to fewer fixes compared to the original notification, where the emails contained links to a HTTPS-enabled Web site. Since browsers allow users to easily determine if the connection is secure, information shared on this *trusted* site might have been more impactful. Hence, we argue that any form of architecture used to disclose vulnerabilities at scale must ensure that trust in the resource can be established regardless of the software used to access it.

All in all, many of the problems discovered in the notification efforts could be addressed in a centralized architecture. Establishing such an infrastructure, which is trusted, can itself easily establish trust in researchers disclosing vulnerabilities, and scales well to a large number of notifications, is key to ensure more successful notifications in the future.

Human Challenges

Next to technical challenges, the previous works highlighted that there are also several human challenges to address.

Sender Reputation and User Distrust— Most notably, when receiving an email from a previously unknown sender, users might show a certain distrust in the message. The number of emails which were received, but discarded due to the recipient's distrust remains unclear, since neither work used, e.g., tracking pixels. Although **(author?)** [1] discussed that the sender of an email did not have a significant impact on the success of a notification of malware-infested sites, we have found evidence to the contrary in our work. More precisely, the German CERT was more inclined to forward our information, simply because they had dealt with us before. In addition, we found that some hosting providers did not react to our notifications to them, but handled the notifications they received from the CERT. Hence, investigating the impact of the sender's reputation remains a viable avenue for future research.

Misunderstood Reports— Another issue are improper reactions due to the recipient misunderstanding the nature of the notification. In our work, we found several instances in which providers threatened to disable accounts for which we had reported vulnerabilities. Moreover, providers also disabled domains of their customers, simply because they misunderstood our message about a vulnerability as an abuse report. Apart from the obvious issue of taking such harsh action on unchecked abuse reports, the underlying issues for these misunderstandings should be studied in further detail. We believe that this again highlights the need to dedicated *security* contacts, rather than only *abuse* contacts.

Subjective Decisions by Intermediaries— The notification campaigns conducted by Li et al. and us showed that using intermediaries such as CERTs often impaired the success of a campaign. While the exact reasons for this are not known, one possible explanation is the priority given to the incoming reports. The problem here, however, is that this prioritization is subjectively done by CERT employees. A vulnerability in WordPress, albeit exploitable in thousands of domains, might not be seen as important considering that the domains do not have a high visitor profile. On the other hand, especially considering the XMLRPC Multicall flaw, which allows an attacker to easily bruteforce username and password combinations [6], might be more critical to lower ranked sites. Under the assumption that the average WordPress user is less aware of security issues, they are more likely to have simple passwords. Hence, an attacker could more easily compromise such sites, e.g., to then host malware on them. We believe that such a problem can, however, only be tackled in conjunction with an architecture which allows disclosure with less human interaction, such that the workload of CERTS can be reduced.

Improving the Fix Rate— Both Li et al. and we observed that even when reports for a vulnerable service were viewed, only about 30-40% of the issues were resolved. Li et al. discussed a number of possible reasons for this, ranging from the notification not reaching the proper party to remediate the flaw to underestimating the impact of a disclosed flaw. Our findings underline these possible reasons. As an example, we found that only 10% of the reports disclosed via the domain registrants eventually lead to a fixed vulnerability for the Client-Side XSS flaws. This might be caused by the fact that the official registrant is a manager rather than a technician and hence might not understand the underlying issue. We therefore feel that investigating factors which led to these relatively low fix rates have to investigated further. For example, the question arises whether the content of a report should depend on the alleged recipient (e.g., the site's security team, a manager, or a regular WordPress user). In addition, other variables such as message length or chosen language should be investigated to determine whether these have significant impact on the fix rate.

Educating Administrators— The previous works have shown that no long-term improvement could be observed as an effect to the notifications. Specifically, we could only observe minuscule differences in the patching behavior of sites that had successfully acted upon our notification when comparing them to the control group. However, in cases where notifications successfully reached the person that can fix the flaw, the previous works missed the opportunity to educate admins about the security of their deployed systems. Hence, another interesting question is whether a notification in conjunction with information on general security best practices can result in a long-term benefit for the affected parties.

Conclusion

In this paper, we mapped the current landscape of large-scale vulnerability notifications. To that end, we covered the lessons learnt from two major papers in this area, which allowed a first glimpse into the problem space. Based on our observations on the problems these works were faced with, we presented a number of follow-up questions, in terms of both technical and human aspects of such notifications. We argue that much research needs to be conducted to understand how the outlined problems can be tackled and that only then notifications can have a positive, long-term effect on the vulnerability ecosystem.

Acknowledgments— This work was supported by the German Ministry for Education and Research (BMBF) through funding for the Center for IT-Security, Privacy and Accountability (CISPA).

References

[1] O. Cetin, M. H. Jhaveri, C. Ganán, M. van Eeten, and T. Moore. Understanding the role of sender reputation in abuse reporting and cleanup. In *Workshop on the Economy of Information Security (WEIS 2015)*.

[2] Z. Durumeric, J. Kasten, D. Adrian, J. A. Halderman, M. Bailey, F. Li, N. Weaver, J. Amann, J. Beekman, M. Payer, and V. Paxson. The matter of Heartbleed. In *IMC*, 2014.

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

[4] F. Li, Z. Durumeric, J. Czyz, M. Karami, D. McCoy, S. Savage, M. Bailey, and V. Paxson. You've got vulnerability: Exploring effective vulnerability notifications. In *USENIX Security Symposium*, 2016.

[5] F. Li, G. Ho, E. Kuan, Y. Niu, L. Ballard, K. Thomas, E. Bursztein, and V. Paxson. Remedying web hijacking: Notification effectiveness and webmaster comprehension. In *WWW*, 2016.

[6] B. Stock, G. Pellegrino, C. Rossow, M. Johns, and M. Backes. Hey, you have a problem: On the feasability of large-scale web vulnerability notification. In *USENIX Security Symposium*, 2016.