Statement from the USENIX Security ’23 Program Committee

This paper was submitted for review by the USENIX Security ’23 Program Committee. The paper received an “Accept Conditional on Major Revision” decision and, following revisions, was ultimately accepted to appear in the Proceedings of the 33rd USENIX Security Symposium. The USENIX Security ’23 Program Committee (PC) and Research Ethics Committee (REC) recognized the technical contributions made in this paper but observed that some of the experiments the authors conducted raise significant ethical concerns. Those concerns resulted in the solicitation of opinions from all non-conflicted members of the REC. Discussion of this paper continued during the PC meeting and afterward. That discussion focused heavily on two ethical concerns arising from the experiments conducted by the authors.

1. Inadvertently collecting potentially sensitive information from the target systems

The experiments carried out by the authors involve running a payload in the exploited target system. This payload collects elements such as file paths and document titles, which were used to gather information about the system and to link responses to requests. The PC noted that these collected items can contain potentially sensitive information. The PC and the REC did not reach a consensus about the authors’ claim that this approach is “minimally invasive.” Some PC and REC members pointed out that the same research goals could have been achieved using less-invasive alternatives. One example suggested during the discussion is to send back a signal from the target system indicating that the attack succeeded, along with information that could facilitate reporting to the server operators before deploying the second stage. An alternative design would be to space out the verification in time so that it would be possible to link responses to requests. Either alternative would allow researchers to contact system operators to report the findings and ask them for additional information about the platform important for understanding and quantifying the impact of the vulnerability.

The PC recognized that the process was approved by the authors’ Institutional Review Board (IRB), which concluded that the potential security improvement achieved by the study outweighs the risks of conducting the experiment. However, the PC noted that the IRB application does not mention the possible involvement of potentially sensitive information in the information collected from target systems.

To address this concern, the PC requested the authors add a thorough discussion about the choice of this experimental setup, the reasons for discarding less-invasive alternatives, and other precautions taken to mitigate risks, including consent processes and data management procedures (e.g., screening procedures for sensitive data and deletion procedures).

2. Conducting tests on live systems absent informed consent

The second ethical concern is related to the exploitation of vulnerabilities in production systems, as conducted by the authors. Despite agreement that measuring and understanding the prevalence and impact of the vulnerability is an important research question, actively seeking to exploit 100k sites, without the operators of those sites being able to opt either in or out, deeply concerned reviewers and the REC. Given these concerns, the PC chairs contacted the authors and asked them for additional clarifications about alternative methods that were considered for identifying vulnerable servers and the reasons for which those methods were rejected. The analysis of the authors’ response led the PC to acknowledge (a) that the response indicates the authors considered the risks and benefits, (b) that the response suggests the authors designed a methodology they intended to minimize harm, (c) that experiments of this nature and scale can pose considerable challenges for informed consent, and (d) that the overall process was approved by the authors’ IRB, which concluded that the potential security improvement achieved by the study outweighs the risks of conducting the experiment.

Nevertheless, sending attack payloads to potentially vulnerable live systems, even when said payloads are intended to be harmless, is an invasive and controversial experimental technique for which the research community does not yet have widely accepted guidelines. During the discussion, some PC and REC members voiced their concerns about the potential for this paper to signal to other researchers that exploiting live systems is a broadly acceptable research practice.

To address this concern, the PC asked the authors to add a discussion about the risks and benefits of this experimental setup.

(continued on next page)

The required changes discussed above did not result in unanimous support for acceptance, and the PC considered alternatives. For example, one suggestion was to ask the authors to repeat the experiments collecting less potentially sensitive information and obtaining consent of site administrators prior to deploying payloads on live systems. Some PC and REC members objected to this option for various reasons, including (a) the justifications for and low risk of the particular experimental setup designed by the authors; (b) the lack of specific community guidelines and best practices about how to conduct this research; (c) the potential to signal to other researchers that a no-exception prior consent requirement exists for experiments on live systems; and (d) the fact that repeating the measurement using a different methodology would not reduce the potential harm of the completed experiment and would introduce new risks.

Ultimately, the PC reached general consensus to accept the paper with this statement describing the decision process and concerns if the above-described requirements were met through a revision. The PC notes that acceptance of this paper is not an endorsement of experimental techniques that actively seek to exploit vulnerabilities in live systems absent informed consent. Any exceptions to informed consent should be carefully justified based on the particular circumstances. The benefits, harms, and risks of various experimental designs should influence the chosen design, including factors like scale and decisions regarding informed consent. An analysis of tradeoffs should occur prior to conducting an experiment as opposed to retrospectively based on findings. The USENIX Security '23 PC hopes that this case serves as a catalyst for thoughtful research processes and community discussion weighing benefits, risks, and possible harms. The PC expects future authors to consider these tradeoffs carefully and address them explicitly.