

I would like to nominate the paper titled "Uninvited Guests: Analyzing the Identity and Behavior of Certificate Transparency Bots" by Kondracki et al. published in USENIX Security 2022. In that paper, the authors described the first distributed honeypot for Certificate Transparency (CT). CT is a relatively recent mechanism where Certificate Authorities must publicly report all the TLS certificates that they issue. While CT is meant to be used for detecting illicitly-issued certificates by malicious/compromised CAs, there have been suspicions that attackers are also tapping into it to identify vulnerable endpoints to attack.

The honeypot system that the authors developed (called CTPOT) issued requests for thousands of certificates over a period of 10 weeks, using unique domains that were never advertised before their experiment. By pointing these domain names to servers running multiple medium-interaction honeypot services, the authors were able to attract attackers abusing the Certificate Transparency program. Their results were eye-opening: their honeypots attracted 1.5 million requests from 32K unique IP addresses that were apparently monitoring Certificate Transparency in search for systems to scan. By using different types of keywords in the domains that they announced, the authors discovered that the domain names that contained trademarks (like paypal or google) and words associated with commonly-vulnerable software (like admin, mysql, etc.) received significantly more requests, compared to their baseline. What I really didn't expect to see in their results was that attackers arrived at their honeypots as early as 12 seconds after they had obtained their TLS certificates (and therefore 12 seconds after their honeypot domains were listed to the Certificate

Transparency log). The authors used TLS fingerprinting to identify the real software that these clients were running (as opposed to what was stated in their user agents) and clustered attackers into campaigns. Overall, their system cleverly combines multiple ideas from cyber security (including honeypots, client fingerprinting, DNS, reverse proxies, and bot detection) to produce the first system capable of accurately differentiating CT-abusing attackers from general attackers who just scan the IPv4/IPv6 space.

Other than the specific abuses that the authors recorded, I find this paper to be particularly important because it shows how a system that was created to benefit the security community (certificate transparency) was converted into a weapon against benign systems, i.e., a way to identify potentially-vulnerable systems that can be compromised. There are many lessons to be learned here to ensure that any future countermeasures that our community builds do not unintentionally provide new offensive capabilities to attackers. Lastly, it is worthy noting that the paper has 3 USENIX artifact badges from USENIX (available, functional, and reproduced) showing how the authors were comfortable having their results evaluated by independent third parties.

For all of these reasons, the "Uninvited Guests" paper has my full support for NSA's best scientific cybersecurity paper competition.