The Cookie Hunter: Automated Black-box Auditing for Web Authentication and Authorization Flaws

Robert Krzysztof Robert Noparlik

October 30, 2023

1 Summary

The proposed system automatically attempts to create accounts on the most popular sites and tests how secure the cookies they set are. It tries to detect if the site supports account creation. If it does, then it attempts to create an account. The system the evaluates if a cookie is vulnerable, for example, by checking if the headers *httpOnly* or *it* attributes are set. If a site is found to be vulnerable, they then deploy a different browser instance to automatically check how much data is leaked.

The paper also goes into detail about how their system's automation and fault tolerance works.

Their findings suggest, that 11K out of 25k domains are vulnerable to eavesdropping attacks and that 23% are susceptible to cookie hijacking through JavaScript.

2 Pros

- The system is robust and supports scanning a large number of websites with features like SSO.
- The system is described in great detail, including alternate execution paths in case of errors.

3 Cons

- Does not support captchas due to higher amount of work and moral dillemas.
- Its robustness makes the system much more complex, especially for such a seemingly simple task (cookie checking).
- Account creation tool not released as OSS.

4 Meaning

One of the most awe-inspiring things about the paper is that the authors tackled the problem of automatic account creation on websites. A simple-looking step required to start analyzing cookies is really hard in practice given the amount of websites on the internet and the prevalence of highly dynamic SSAs, which make scraping harder.

Although not the first study to automatically check for cookie vulnerabilities, the authors improved on the previous method to detect AuthCookies. Once detected, they can be audited and used to check privacy leakage.

5 Discussion and Questions about the paper

How dangerous are those cookie security issues actually, especially if a lot of modern sites might not serve HTTP content at all?