Reducing Vulnerability Information Leakage from Public IDS Rules

Nathan Qiu[1], Yigitcan Kaya[2]

1. University of California, Santa Barbara

UC SANTA BARBARA

1. Background

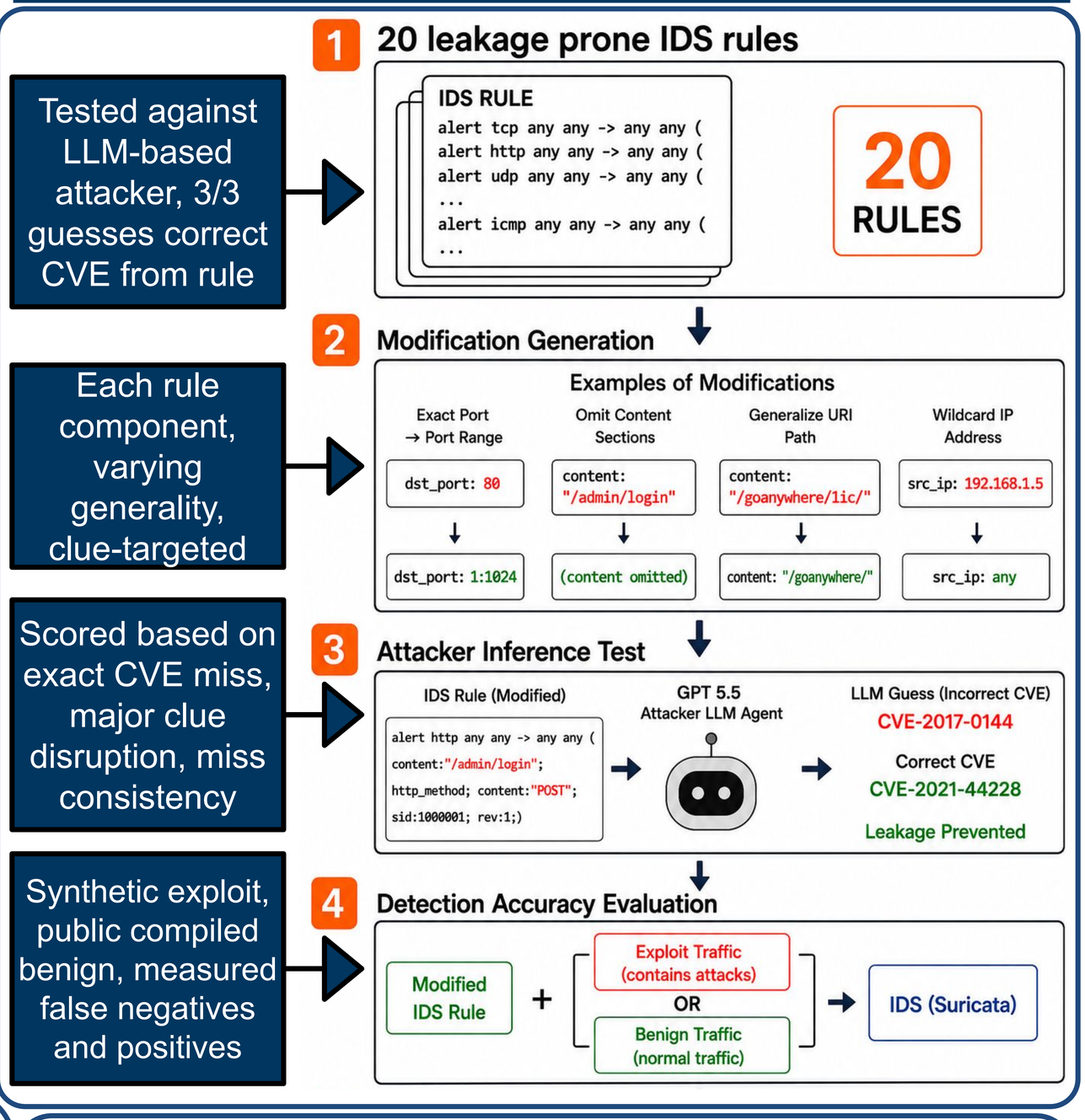
- Intrusion detection systems (IDS) use signature-based rules to detect known cyberattacks. Rules are released to the public to allow users to protect their systems
- Attackers can infer the vulnerability of a system from the public rules
- **Can we modify IDS rules to reduce vulnerability information leakage without sacrificing attack detection?**

```
alert http any any -> $HOME_NET any (
  flow:established,to_server;
  http.method; content:"POST";
  http.uri; content:"/goanywhere/lic/accept?bundle=";
  fast_pattern; startswith;
  sid:2044143; rev:1;
)
```

Example IDS rule for the Fortra GoAnywhere MFT Pre-Authentication Remote Code Execution Vulnerability

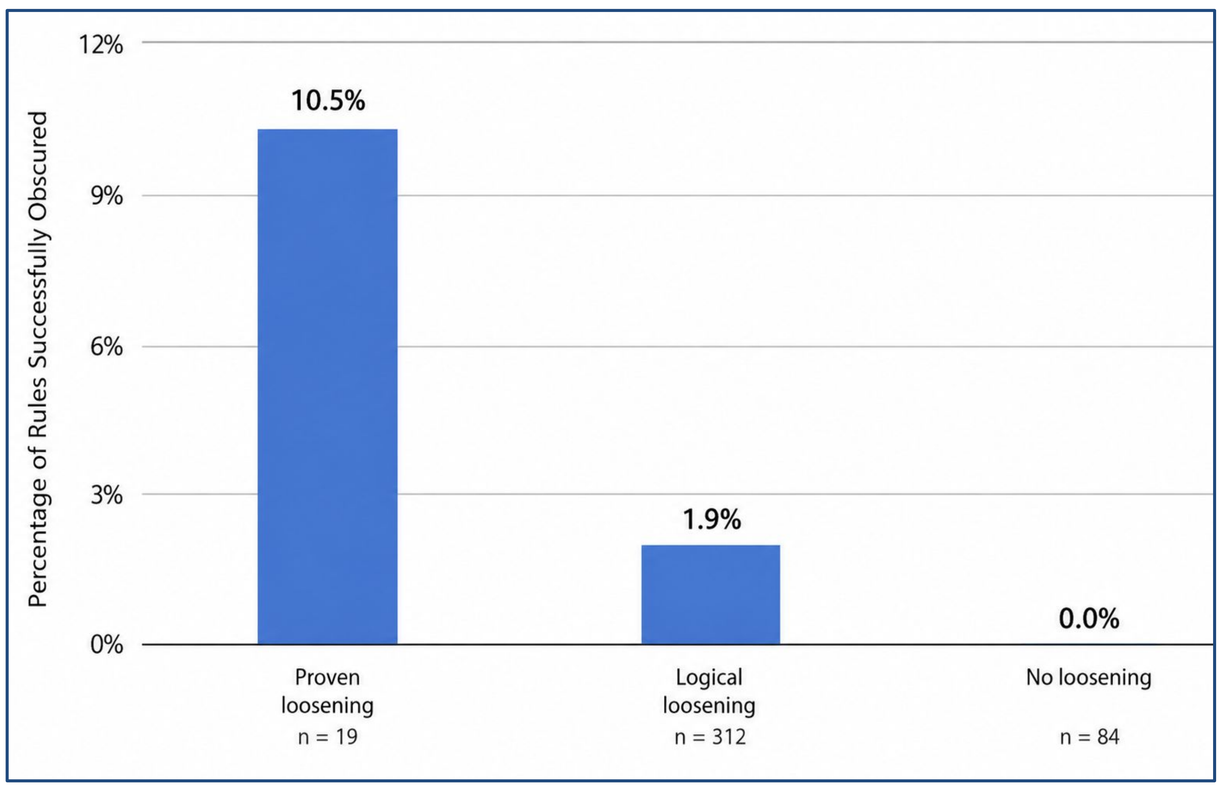
Dangerous clue including the product name explicitly, attacker could easy infer vulnerability

2. Experimental Design



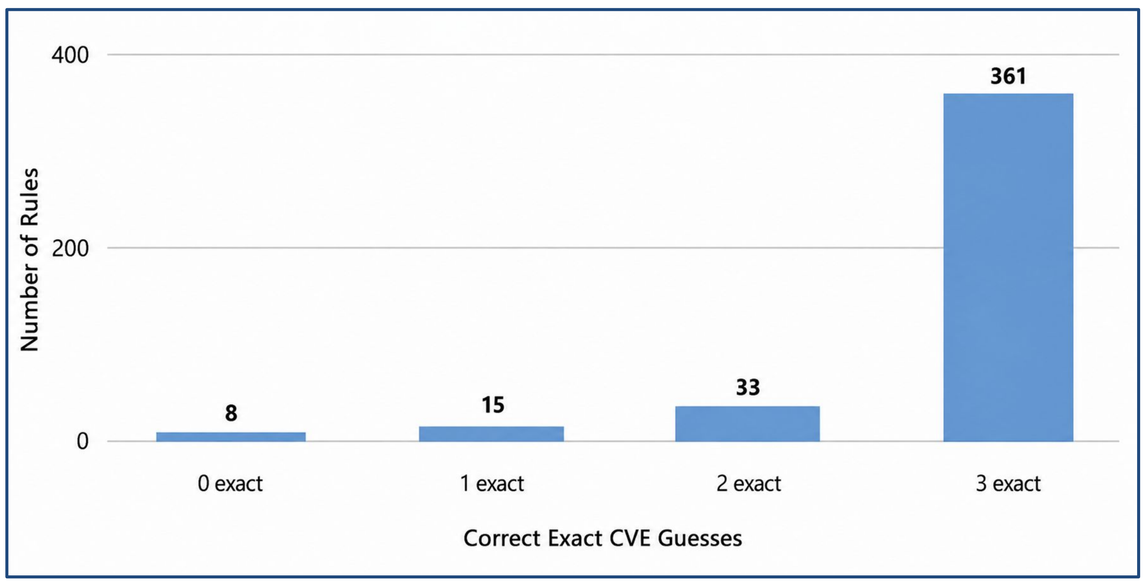
3. Experiment Results

Trade off between rule obscurity and rule looseness(detection precision)



- Out of 417 modified rules, 19 triggered false positives in benign traffic tests of which 2 (10.5%) made the attacker miss 3/3 times.
- 312 modified rules triggered no false positives, but logically loosened the traffic detection of the rule, of which 6 (1.9%) made the attacker miss 3/3 times
- 84 modified rules only changed the rule representatively, with the logical detection of the rule staying exactly the same leading to 0 3/3 attacker miss cases

Three attacker trials separate stable effects from one-shot noise



- 8/417 modified rules (1.9%) obscured the rule substantially, making the attacker-LLM miss 3/3 trials.

4. Conclusions and Future Directions

Conclusions

- Single-modifications rarely obscure rules significantly
- Successfully obscured rules come with a tradeoff of having to sacrifice traffic detection precision

Improvements

- Experiment with combinations of modifications
- Increase rule dataset and increase attacker trials
- Compile more representative benign traffic tests

AI Institute for Agent-based Cyber Threat Intelligence and Operation

Supported by the National Science Foundation under grant no. IIS-2229876 and is supported in part by funds provided by the NSF, by the Department of Homeland Security, and by IBM.

https://action.ucsb.edu