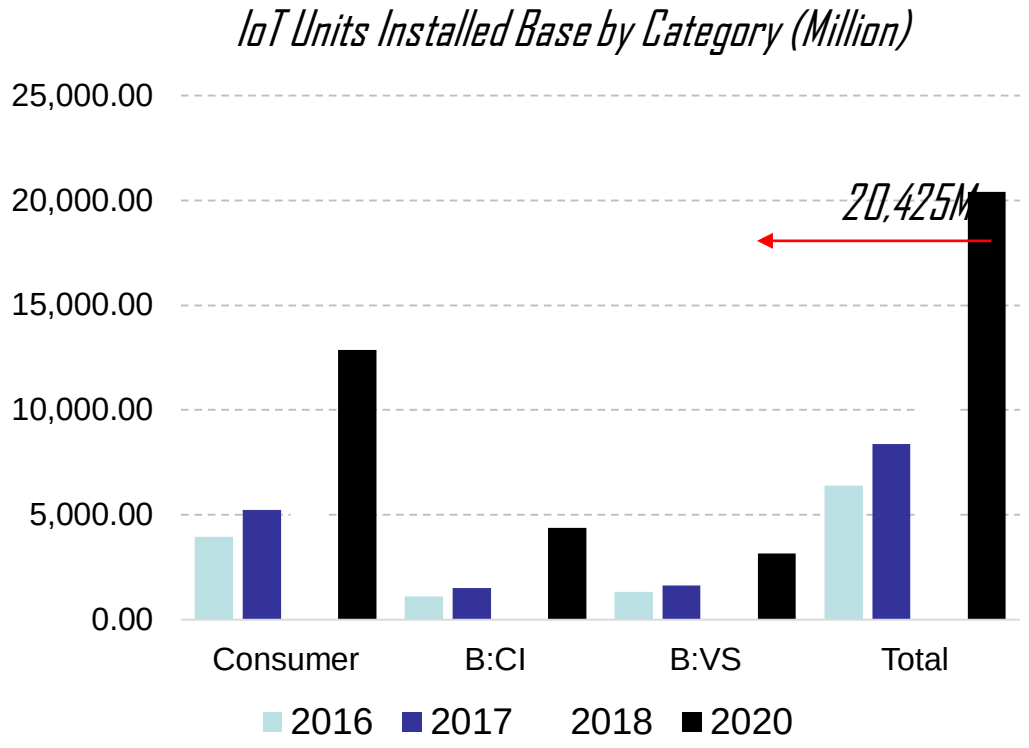
Understanding and Securing Device Vulnerabilities through Automated Bug Report Analysis

Xuan Feng, Xiaojing Liao, XiaoFeng Wang, Haining Wang, Qiang Li, Kai Yang, Hongsong Zhu, Limin Sun

USENIX Security 2019

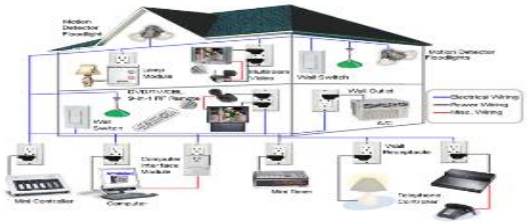
Internet-of-Things (IoT) Devices

Various IoT devices connected to the Internet



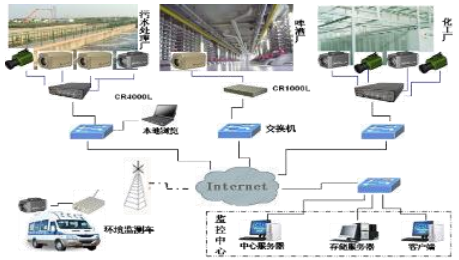
5.5 million new IoT devices every day
20 billion by 2020 (*By Garnter*)

IoT devices yield substantial security challenges

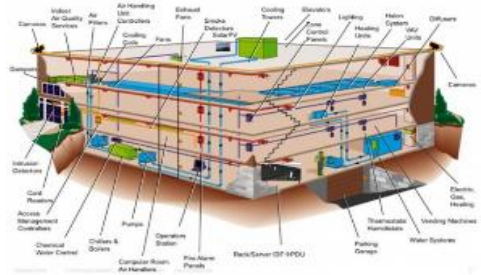


Smart Home

Wearable computing



Surveillance

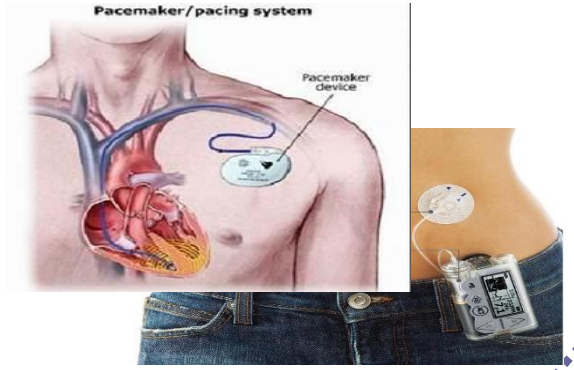


Smart Building

Smart Grid

Urban water/gas

IoT Security Concerns



Barnaby Jack hack wireless Pacemaker

2016 DDoS attacks Dyn Service

2010 BlackHat Jackpotting hack ATM

Australia SCADA
sewage into the river and coastal waters

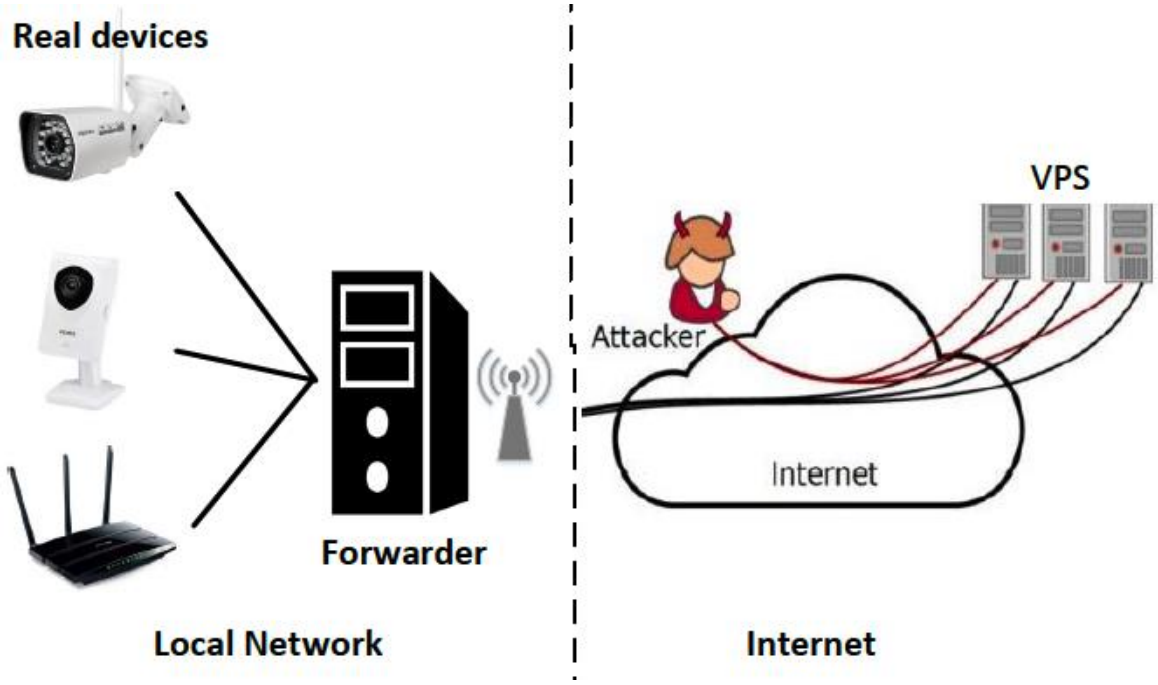
Know yourself and know your enemy, and you will never be defeated.

- ***Sunzi's*** **Art of War** 孙子兵法

Understanding the perilous IoT world.

- Real device honeypot.
  - VPS as relay hosts
  - reverse SSH tunneling

- Simulated Honeypot
  - whose default configurations (such as default page and HTTP response header/body) have been modified to simulate real devices.



The infrastructure of real device honeypot

Understanding the perilous IoT world.

- From May to July in 2018, our honeypots gathered 190,380 HTTP requests from 47,089 IPs across 175 countries.

	Real devices	Simulated honeypots
Malicious (Targeted)	20	~300
Malicious (Blind-scanned)	121	~1,560
Benign	11,451	176,764
Unknown	10	~154
Total	11,602	178,778

Traffic analysis of deployed honeypots.

- More than 90% of malicious attacks exploit the *known* vulnerabilities.

Understanding the perilous IoT world.

Name	Vulnerabilities
IPCAM exploits	Pre-Auth Info Leak
Huawei Exploits	Command Execution
iotNigger	Netis Backdoor
Brickerbot	More than 30 vulnerabilities

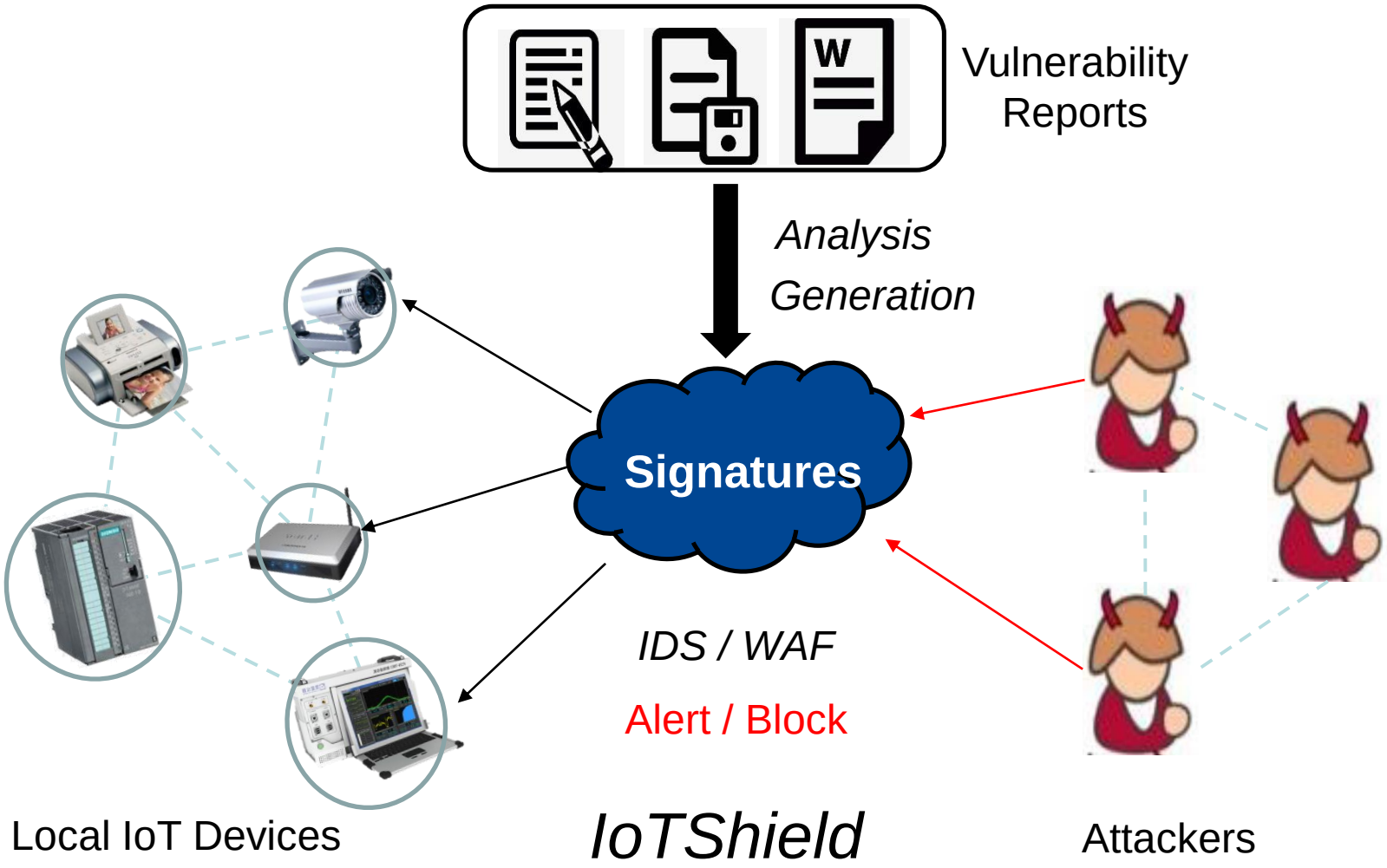
Underground IoT attack tools

Name	Vulnerabilities	Year
IOT Reaper [24]	10 vulnerabilities	2017
Hajime [23]	at least 3 vulnerabilities	2016
Satori [33]	2 vulnerabilities	2018
Brickerbot [5]	21 vulnerabilities	2017
Masuta [25]	bypass & command execution	2018
Amnesia [2]	remote code execution	2017

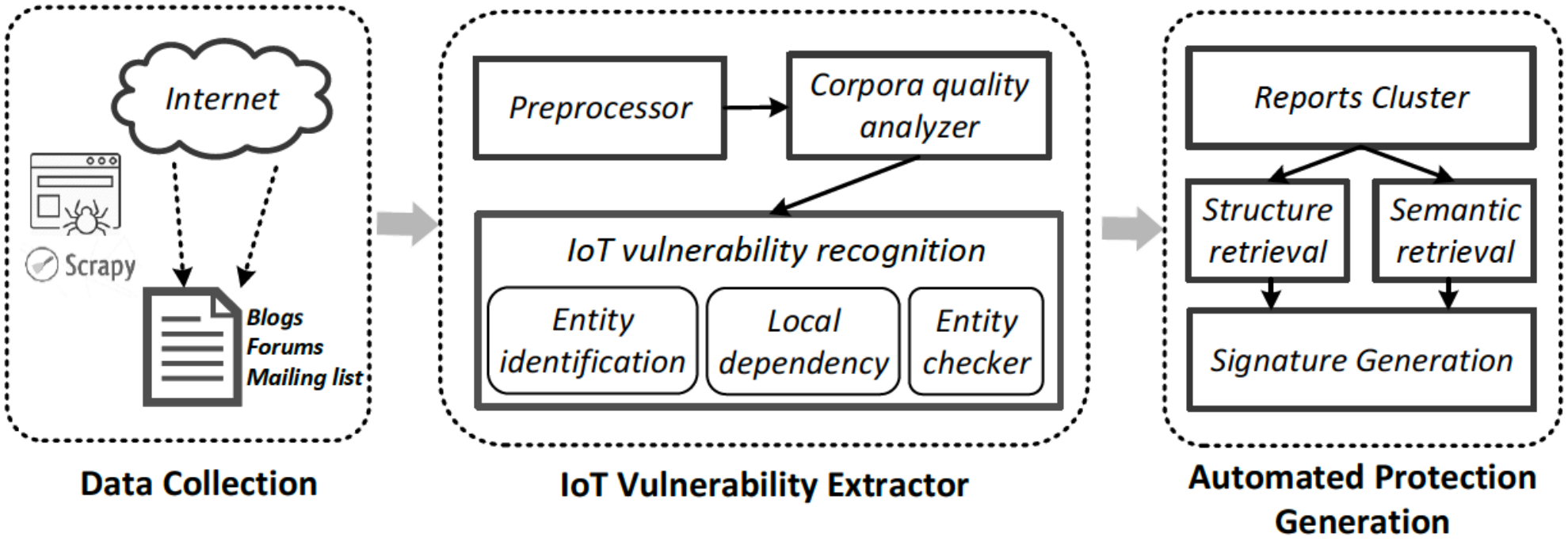
Known IoT attack activities

- To validate the findings made from the honeypots, we further analyzed four underground attack toolkits and six well-documented IoT botnets.
- The exploitation of the *known* vulnerabilities also exists in underground attack toolkits and known IoT attack activities.

Automated Signature Generation

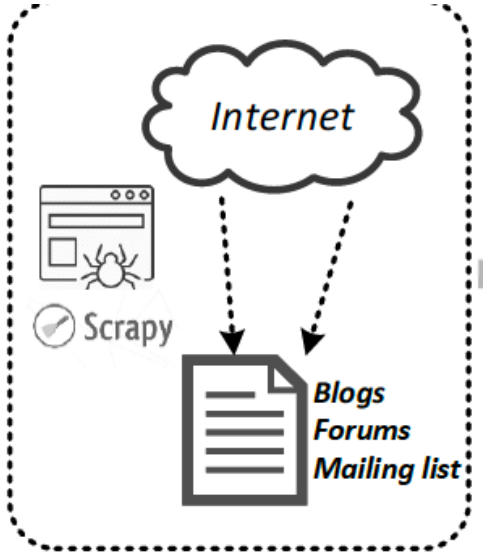


Automated Signature Generation



IoTShield

Data Collection



Data Collection

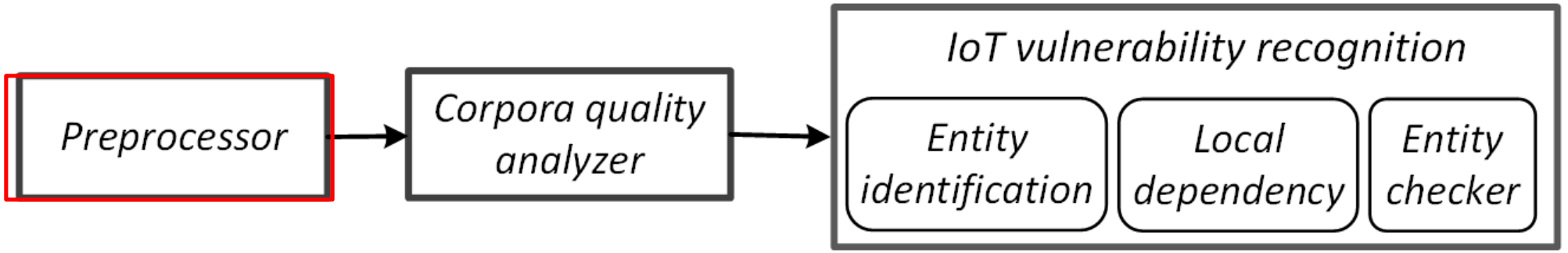
wget

scrapy

Categories	Website	Reports	IoT reports
Personal	s3cur1ty.de/advisories	28	16
Blogs	pierrekim.github.io	18	13
	gulftech.org	129	5
Forums	seclists.org/fulldisclosure	108,647	1,219
Team	coresecurity.com	390	31
Blogs	vulnerabilitylab.com	2,122	39
	blogs.securiteam.com	1,925	42
Mailing lists	seclists.org/bugtraq	85,593	1,591
Data	exploit-db.com	39,380	895
Archive	packetstormsecurity.com	97,093	1,951
	0day.today	30,177	834
	seebug.com	56,413	690
	myhack58	7,311	150
Total	-	42,9795	7,514

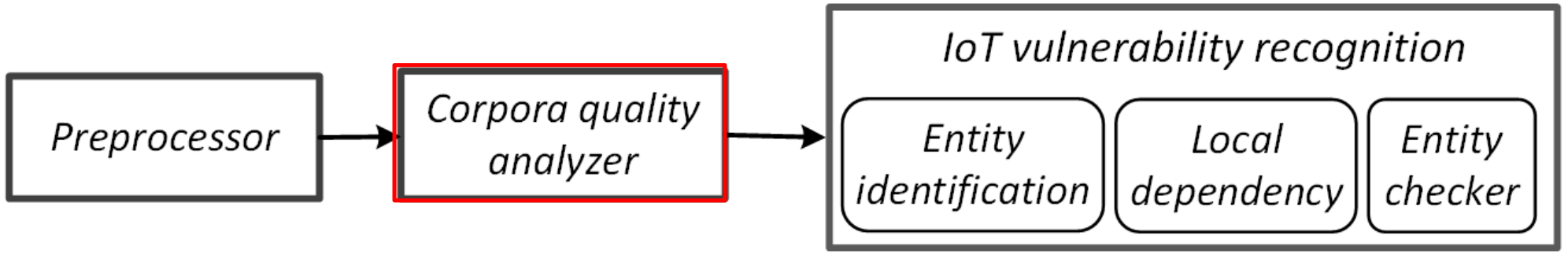
List of vulnerability reporting websites

IoT Vulnerability Extractor



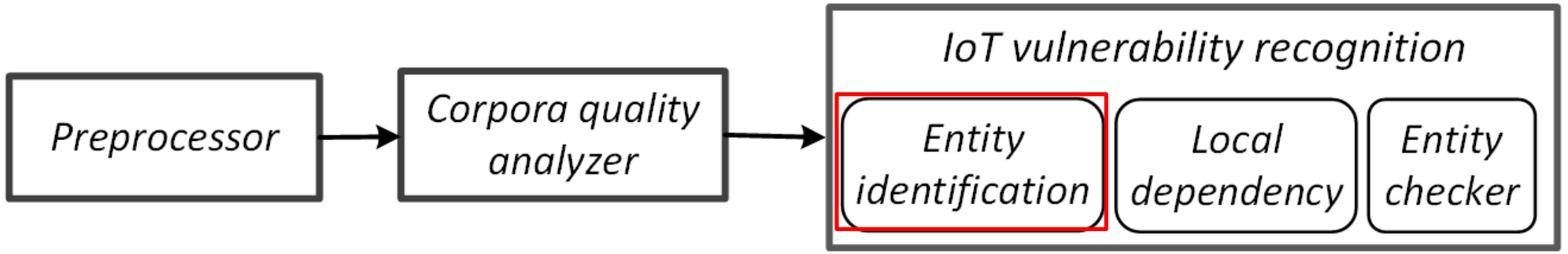
- Remove the textual information irrelevant to vulnerabilities documents
  - ✓ such as advertisements, pictures, dynamical scripts, and navigation bar
- Keep URLs, document titles, authors, and publication dates.

IoT Vulnerability Extractor



- Remove the textual information irrelevant to vulnerabilities documents
  - ✓ The percentage of dictionary words (82%)
  - ✓ The number of hyperlinks (25 hyperlinks)
- Performance of these two heuristics
  - ✓ 100 documents being filtered.
  - ✓ 0% false positives

IoT Vulnerability Extractor

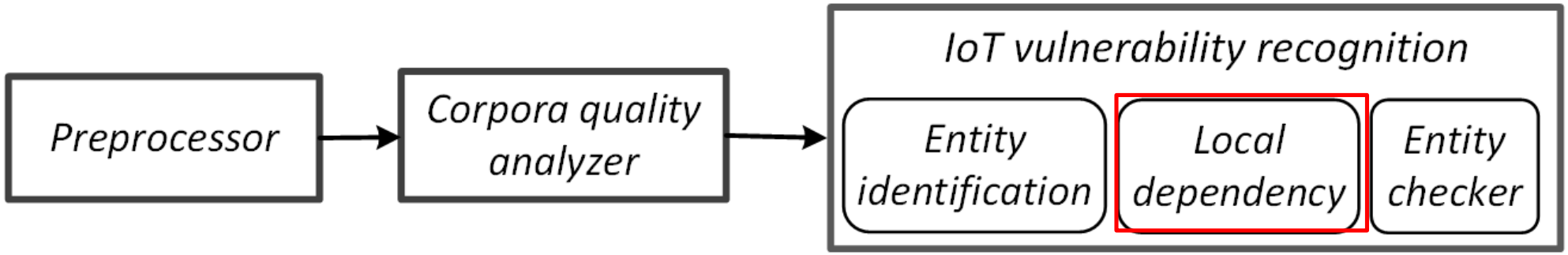


- To identify these individual entities, we utilized keyword and regular expression based matching.
  - corpus-based: device types, vendor names and vulnerability type
  - rule-based: use regular expressions to extract the product name entity.

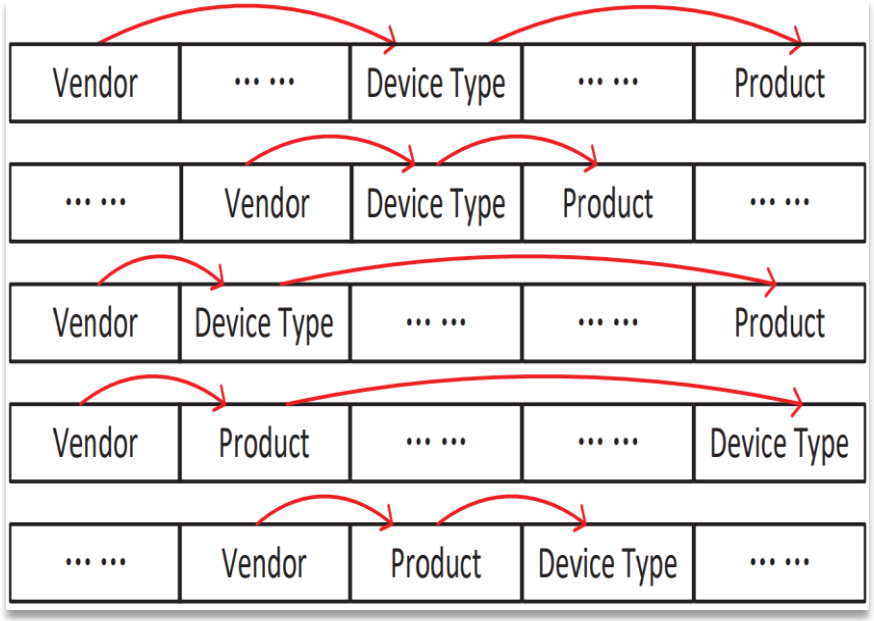
Entity	Context terms
Device Type	camera, ipcam, netcam, cam, dvr, router nvr, nvs, video server, video encoder, video recorder diskstation, rackstation, printer, copier, scanner switches, modem, switch, gateway, access point
Vendor	1,552 vendor names
Product	[A-Za-z]+[-]?[A-Za-z!]*[0-9]+[-]?[-]?[A-Za-z0-9] *∧[0-9]2,4[A-Z]+
Vuln type	733 CWE, 88 abbreviations
Version	(?:version[:.]*([\w-][\w.-]+) ve?r?s?i?o?n?s?[:.]*([\d-][\w.-]+)
CVE	CVE-[0-9]{4}-[0-9]{4,}

Context textual terms

IoT Vulnerability Extractor

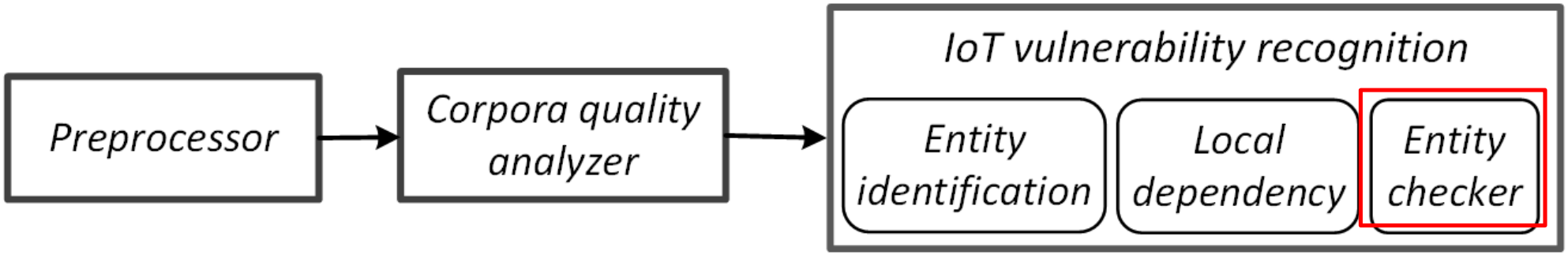


- Poor performance :
  - high FGs in device type/product name.
  - irrelevant webpages include keywords of device type such as “switch”.
  - a phrase that meets the requirement of regex for a product name.
- True IoT entities always have strong dependence upon one another.
  - D-Link DIR-600 or Foscam IPcamera



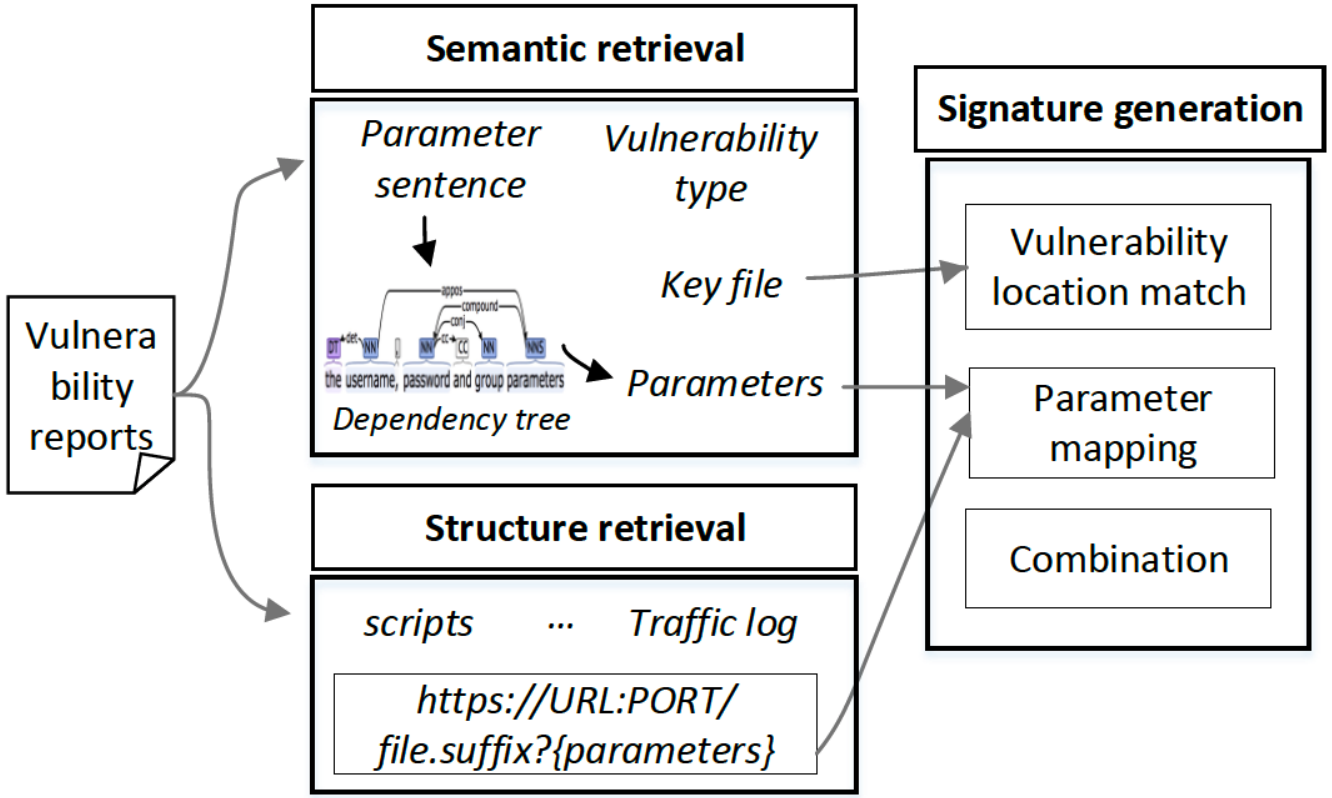
The local dependency of the device entity

IoT Vulnerability Extractor



- Entity checker
  - Search extracted entities (e.g., D-Link DIR-600) in Google
  - Calculate the cosine similarity between the extracted entities and the title of the search results
  - If the similarity is extremely low (e.g., 0.08), the extracted entity is classified as non-IoT

Automated Protection Generation



The architecture of signature generation.

Examples - Automated Protection Generation

13) Authenticated command injection in PwdGrp.cgi The PwdGrp.cgi uses the username, password and group parameters in a new user creation or modification request in a system command without validation or sanitization. Thus and attacker can execute arbitrary system commands with root privileges. We are aware that this vulnerability is being exploited in the wild!

Traffic log:
GET cgi-bin/supervisor/PwdGrp.cgi?action=add&user=test&pwd=;reboot;&grp=SUPERVISOR&lifetime=5%20MIN HTTP/1.1
Host: 107.xx.8.xx
Connection: keep-alive
Accept-Encoding: gzip, deflate
Accept: */*
User-Agent: python-requests/2.18.4

Vulnerability Type: command injection
Vulnerability file: PwdGrp.cgi
Vulnerability parameters: username, password, group

Genera format: http://< *DEVICE_IP* >/cgi-bin/supervisor/PwdGrp.cgi?action=add&user=test&pwd=;reboot;&grp=SUPERVISOR&lifetime=5%20MIN

Vulnerability-based signature

http://<DEVICE_IP>/cgi-bin/supervisor/PwdGrp.cgi?action=add&user={command}&pwd={command}&grp={command}&lifetime=5%20MIN

Snort format signature

alert tcp any any -> any $HTTP_PORTS (content:"/cgi-bin/supervisor /PwdGrp.cgi"; http_uri; pcre:"/[?&](user|pwd|grp)=[^&]?([\x60\x3b\x7c]|echo|ping|cat|reboot|\x3c\x3e\x24]\x28|%60|%20|%3b|%7c|%26|%3c%28|%3e%28|%24%28)/iU";)*

Evaluation - Vulnerability extractor

- We randomly sampled 200 reports from those identified for manual validation and achieve a precision of 94%.
- In total, we collected 7,514 IoT vulnerability reports from 0.43 million articles. These reports disclose 12,286 IoT vulnerabilities, with roughly 1.6 each on average.

Device Vendor	Num	Device Type	Num
Cisco	1,264	router	3,700
D-Link	988	switch	1,422
Linksys	539	camera	1,248
Netgear	522	firewall	1,101
HP	485	gateway	1,032
Symantec	299	modem	843
TP-Link	255	access point	478
Zyxel	229	printer	408
Huawei	195	nas	338
Asus	180	scanner	176

Top 10 vendors and device types of affected devices.

	Vulnerability type	Num
1	Denial of service	975
2	CSRF	902
3	Buffer overflow	869
4	Command injection	806
5	XSS	775
6	Authentication bypass	763
7	Command execution	458
8	Information disclosure	407
9	Directory traversal	307
10	Privilege escalation	276

Top 10 vulnerability types.

Evaluation - Rule generation effectiveness

- 190K HTTP requests collected from real IoT devices and honeypots
  - ✓ simulators: 178,778 HTTP requests related to 141 attack; 26 unique attack scripts; the rest is benign traffic.
  - ✓ real-device honeypots: 11,602 HTTP requests in 1,860 attacks generated by 81 unique attack scripts.
- Macbook Pro with 2.6GHz Intel Core i7 and 16GB of memory.

Dataset	Precision	Recall	False Positive Rate
Real devices	97%	83%	0.01%
Honeypot	98%	93%	0.06%

- Long-time (1 year) traffic captured in an industrial control system HMI honeypot 7,396 alerts of exploiting the HMI system. After manually checking the
  - ✓ 7,396 alerts, we confirmed that about 6,705 alerts were indeed IoT attacks.
  - ✓ The rest of the alerts were confirmed to have attacked other vulnerabilities on common web servers.

Performance

Signature generation

Stage	Running time (s)	Percentage
Data collection	0.386	51%
IoT vulnerability extractor	0.154	21%
Rule generation	0.210	28%
Overall	0.750	100%

Running time at different stages. Time cost of IoTShield for automatic rule generation is low in practice

Rule inspection

- Two-hour real-world traffic captured on the edge router of a research institution (53G)
- IoTShield induces little overhead to IDS

***without* IoTShield**	***with* IoTShield**
426.28s	+0.13s

Conclusion

- *New discovery*
  - IoT vulnerabilities are publicly available and easy to exploit, and today's IoT attacks almost exclusively use known vulnerabilities for mounting malicious attacks.
- *New defense*
  - Our findings lead to the design of IoTShield, a simple yet effective IoT vulnerability-specific signature generation system for intrusion detection systems, which significantly raises the bar for IoT attacks.

Thank you!

Q&A