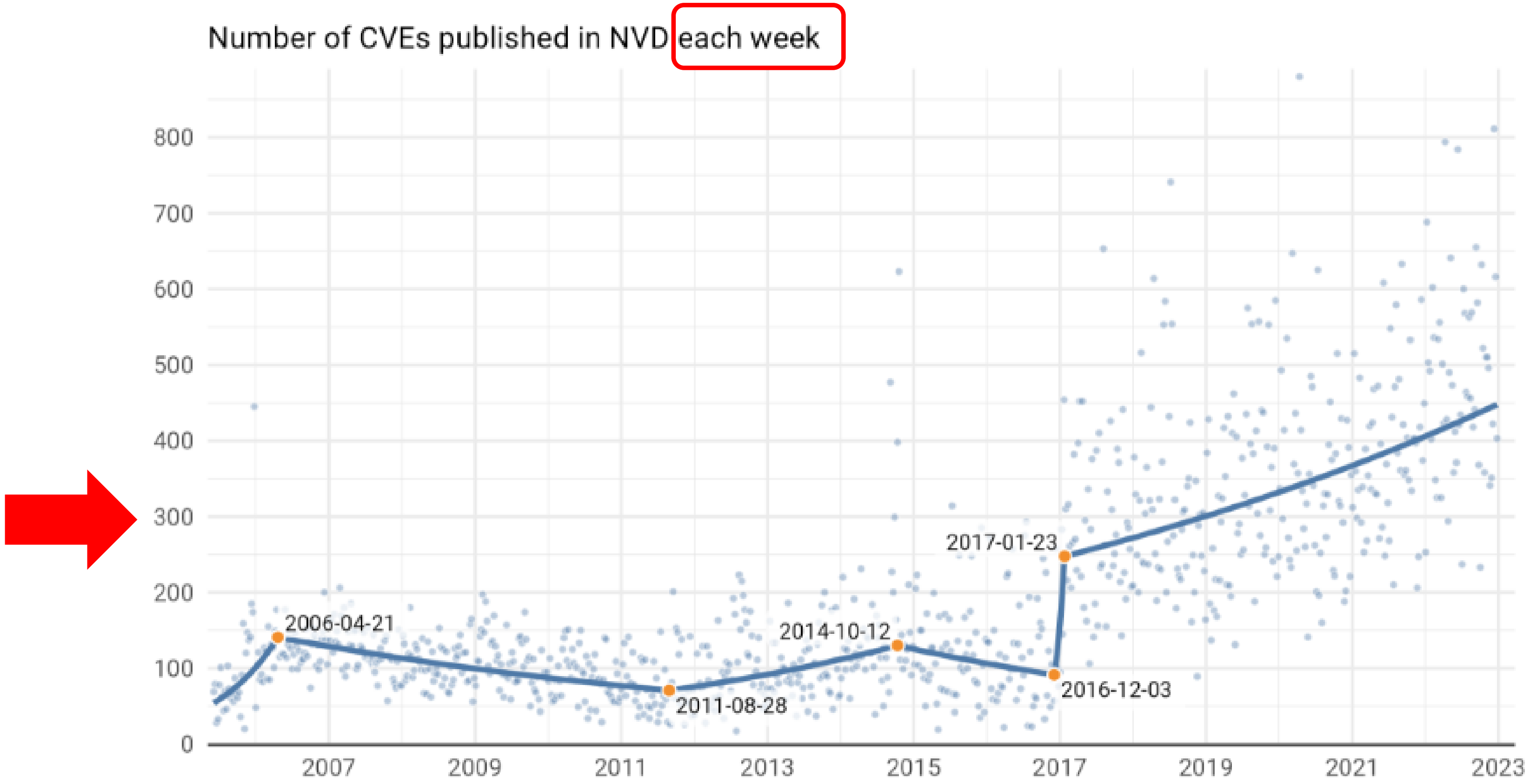
Vulnerability Prioritization

New CVEs


Number of CVEs published in NVD each week
800
700
600
500
400
300
200
100
0
2007
2009
2011
2013
2015
2017
2019
2021
2023
2006-04-21
2011-08-28
2014-10-12
2016-12-03
2017-01-23

Facts about Patch Application (III)

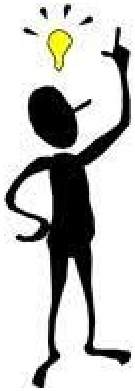
~~4. Owner/Admin applies patch~~

- …
- **There are just too many vulns that might need patching**
- **Fundamental problem**

Key idea

- **Basic fact:**
    - **Very few CVEs are actually exploited**
    - Just to have an idea: **≈5%** of all CVEs **(!)**

- Focus only on those CVE



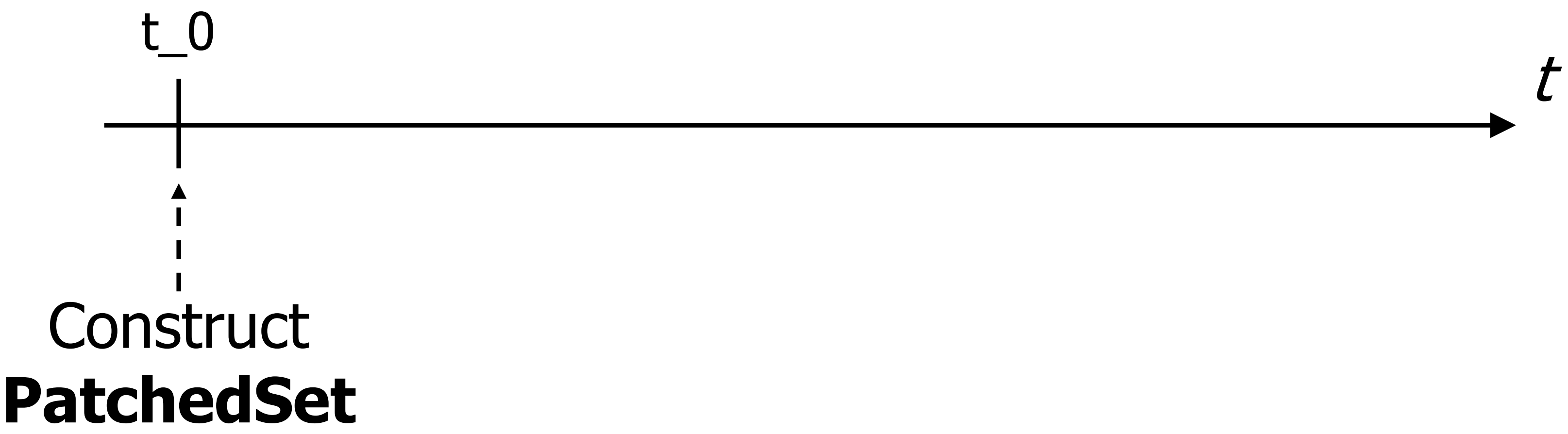
Fundamental problem

- **Basic fact:**
  - **Very few CVEs are actually exploited**
  - Just to have an idea: **≈5%** of all CVEs **(!)**

- Predicting which CVEs will be indeed exploited is **very difficult**

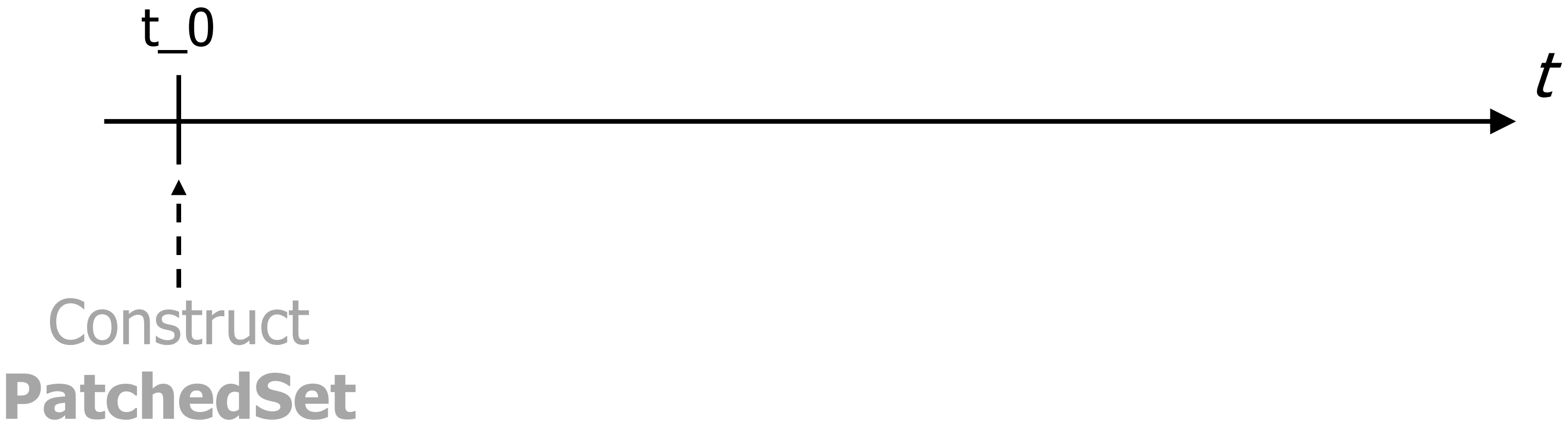
Exploit Prediction: Problem Definition

Exploit Prediction: Problem Definition (I)



- We want to define a criterion for choosing **which vulnerabilities to patch**
- Subset of **all known vulns** at t_0
  - An organization should focus only on vulns on its systems (and their risk)

Many possible criteria



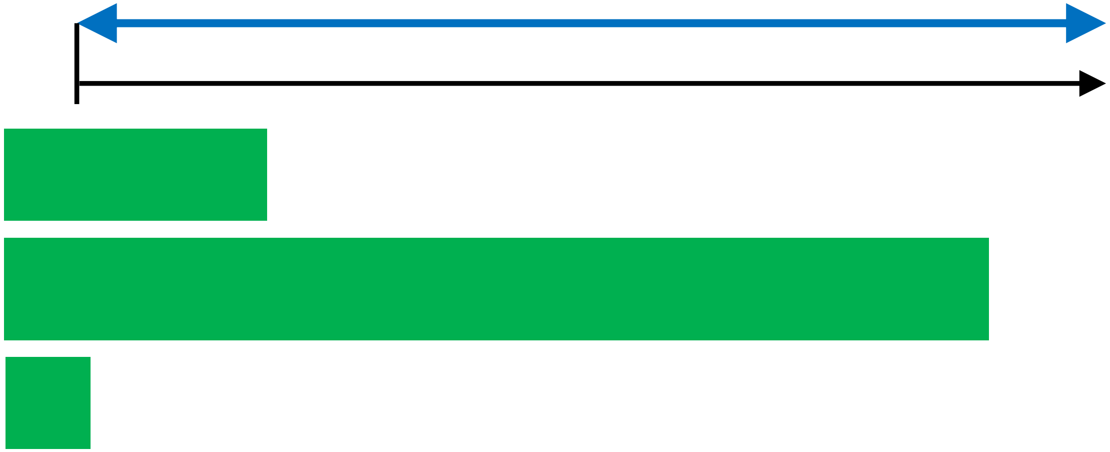
t_0

t

Construct **PatchedSet**

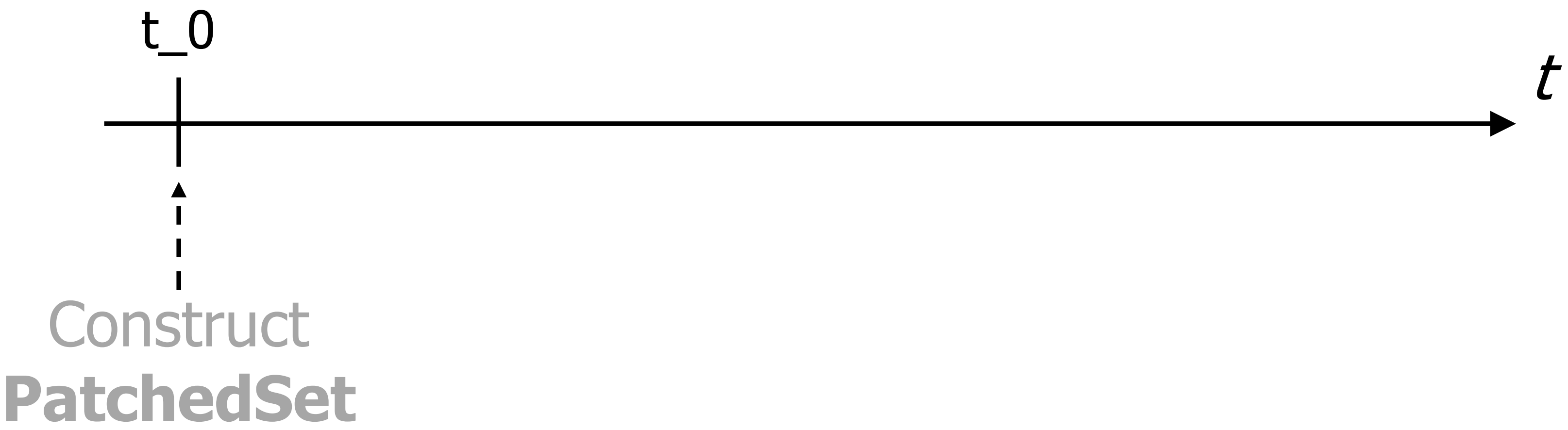
- All vulns with CVSS Critical
- All vulns with remote injection
- All vulns of Windows software
- …

Patching Effort



- Size of PatchedSet = **Patching Effort**
- It depends on the **criterion** used
  - All vulns with CVSS Critical
  - All vulns with remote injection
  - All vulns of Windows software
  - …

Hhmmm...


t_0

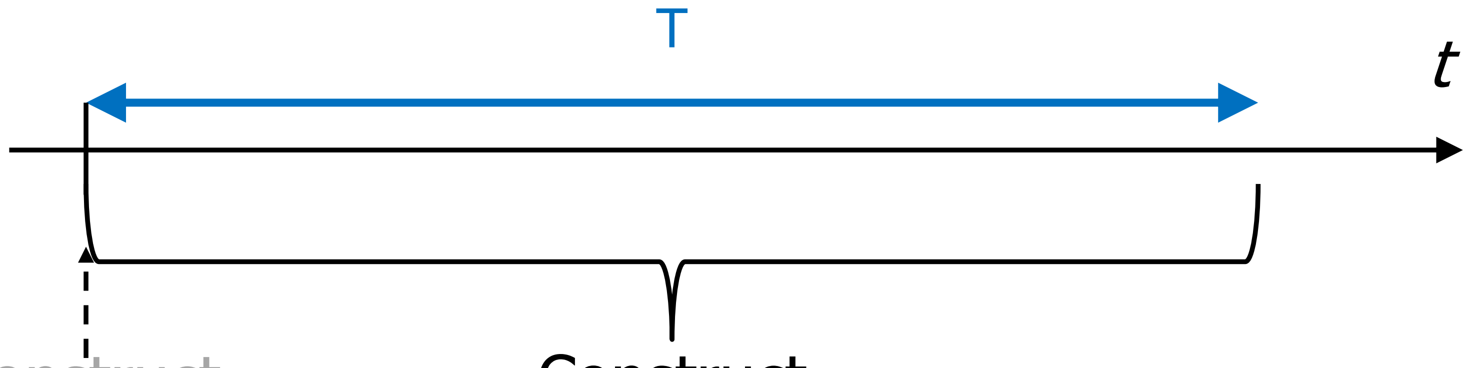
t

Construct

PatchedSet

How to assess a given criterion?

Exploit Prediction: Problem Definition (II)

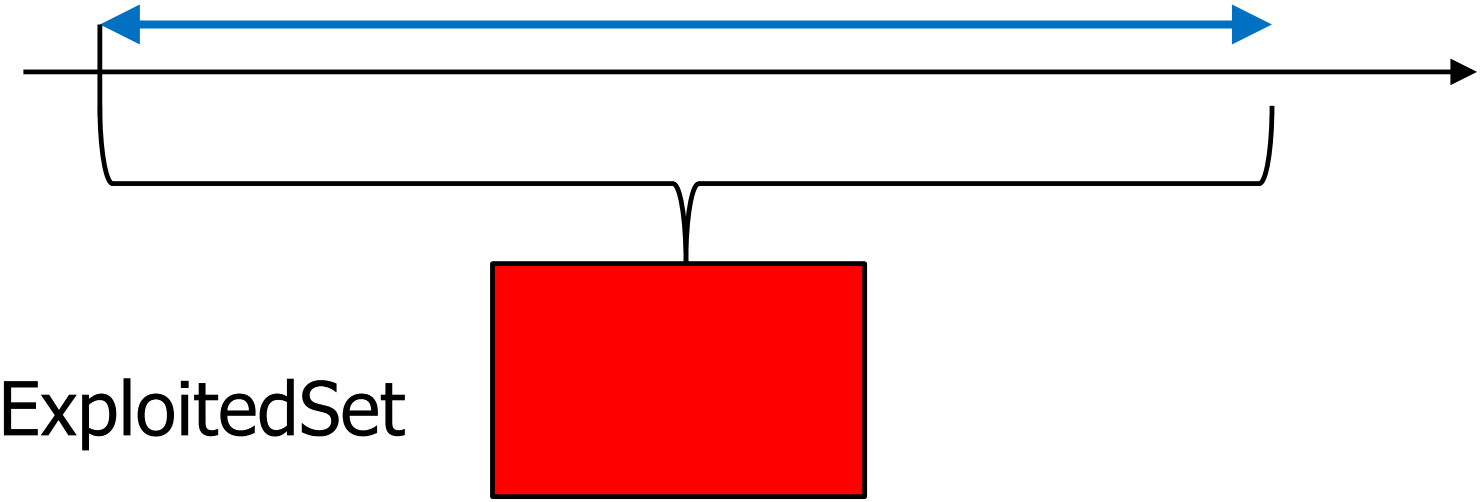


Construct **PatchedSet**

Construct **ExploitedSet**

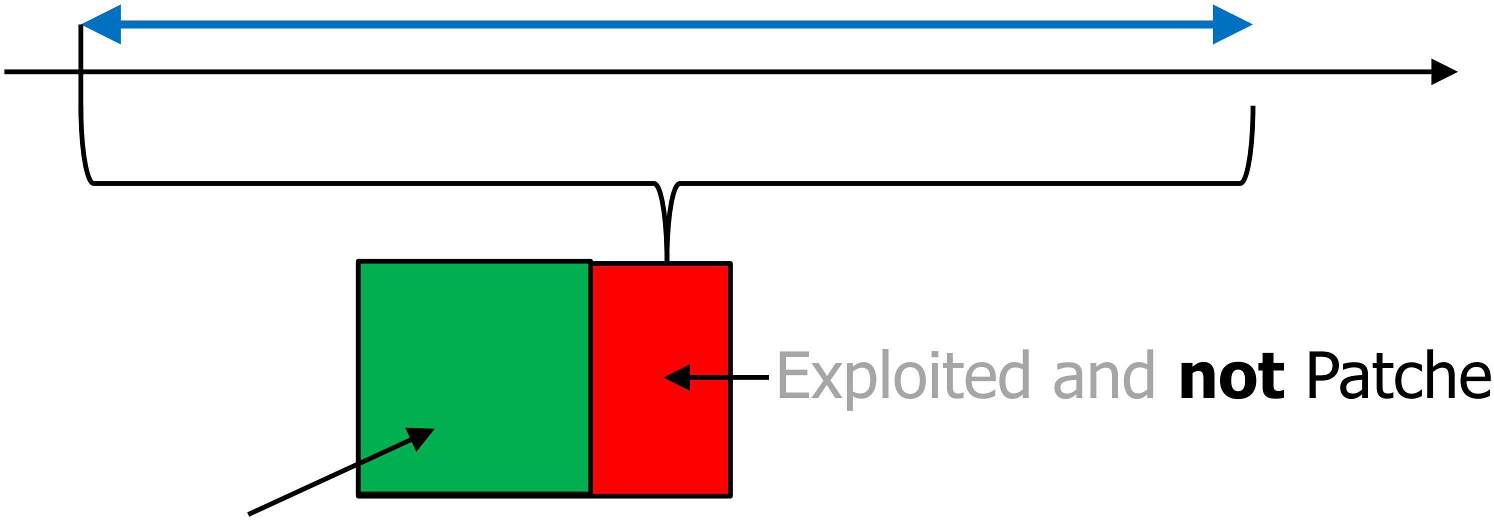
1. We observe which vulnerabilities have been **actually exploited worldwide**
   - Approximation by collecting many intelligence feeds
2. We "compare" PatchedSet and ExploitedSet

Coverage (≈Recall) (I)



How many have been **patched**?

Coverage (≈Recall) (II)



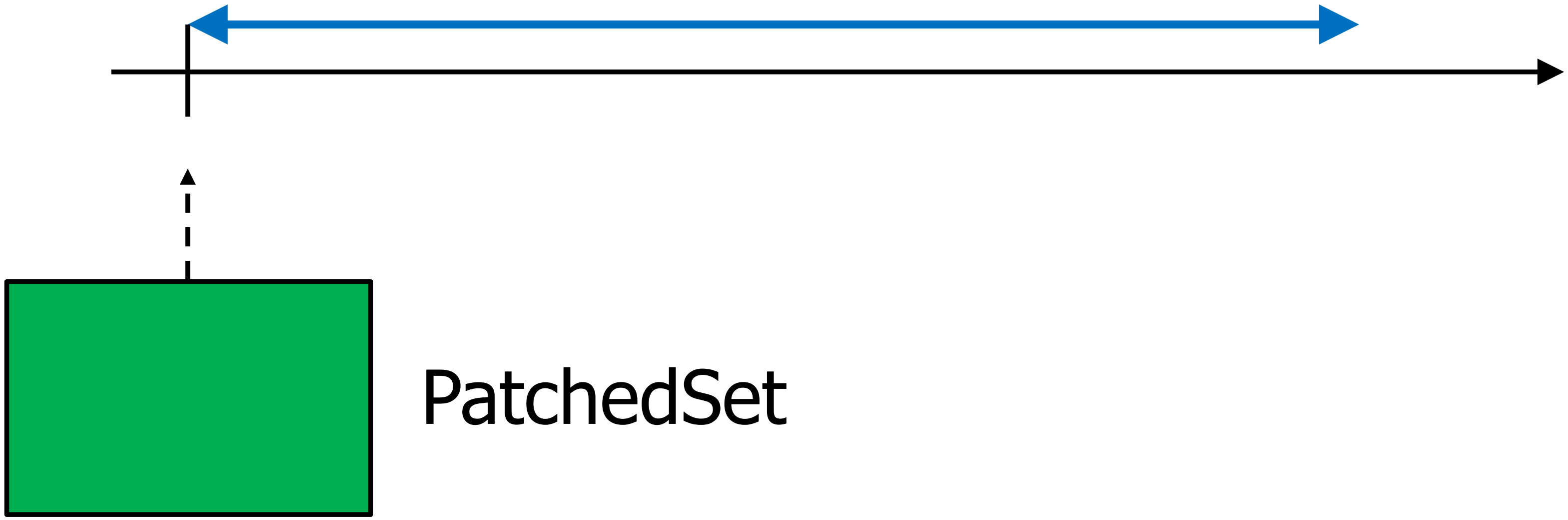
Exploited and **not** Patched

Exploited and Patched

Coverage

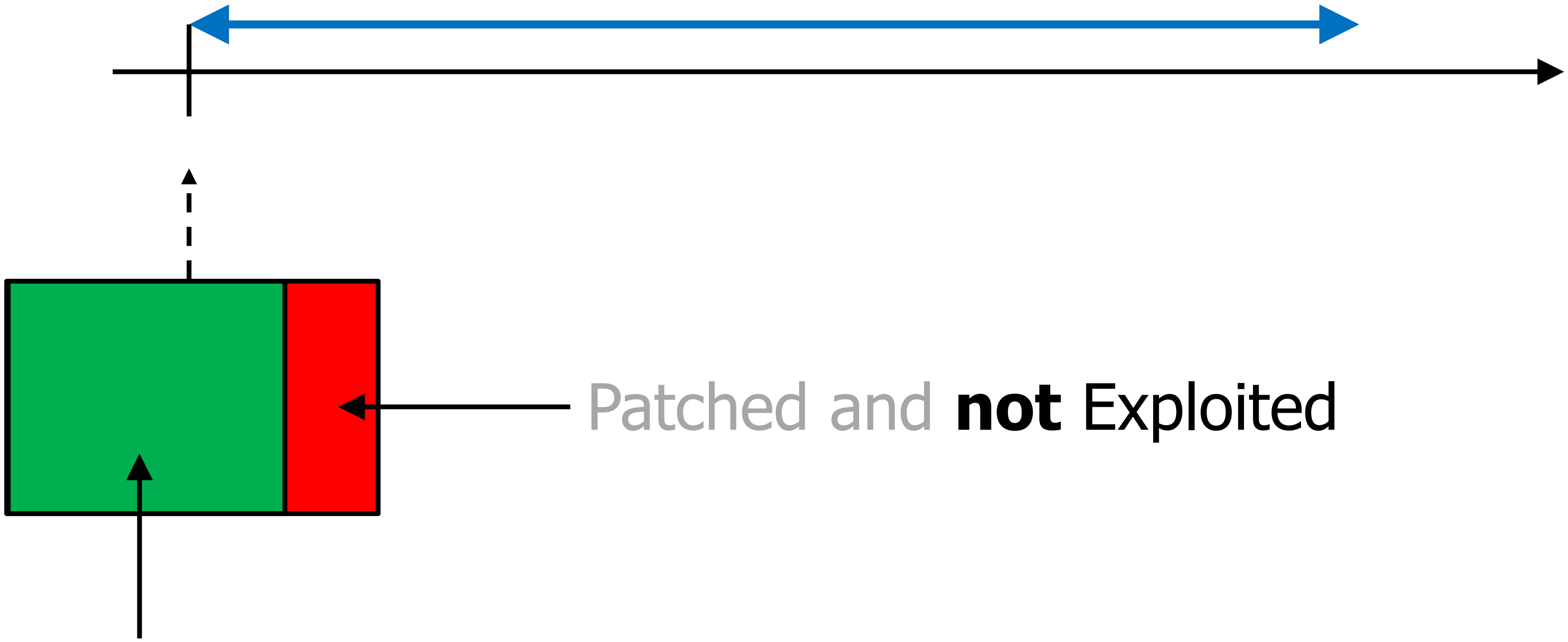
$$\frac{\#(\text{Exploited and Patched})}{\#(\text{Exploited})}$$

Efficiency (≈Precision) (I)



How many have been **exploited**?

Efficiency (≈Precision) (II)



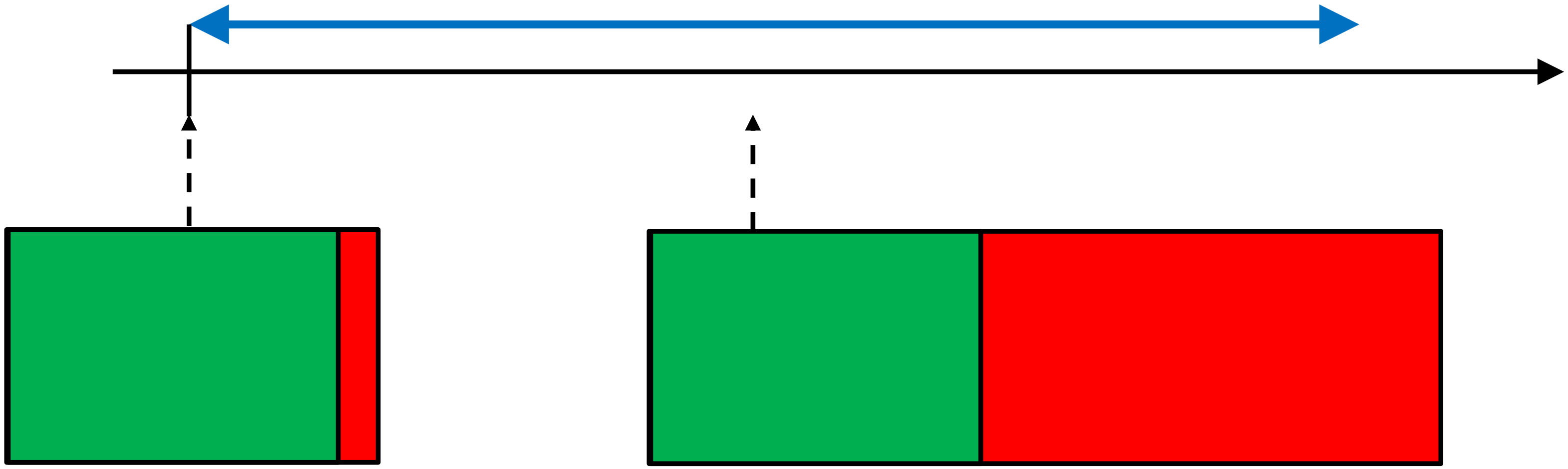
Patched and **not** Exploited

Patched and Exploited

Efficiency

$$\frac{\#(\text{Patched and Exploited})}{\#(\text{Patched})}$$

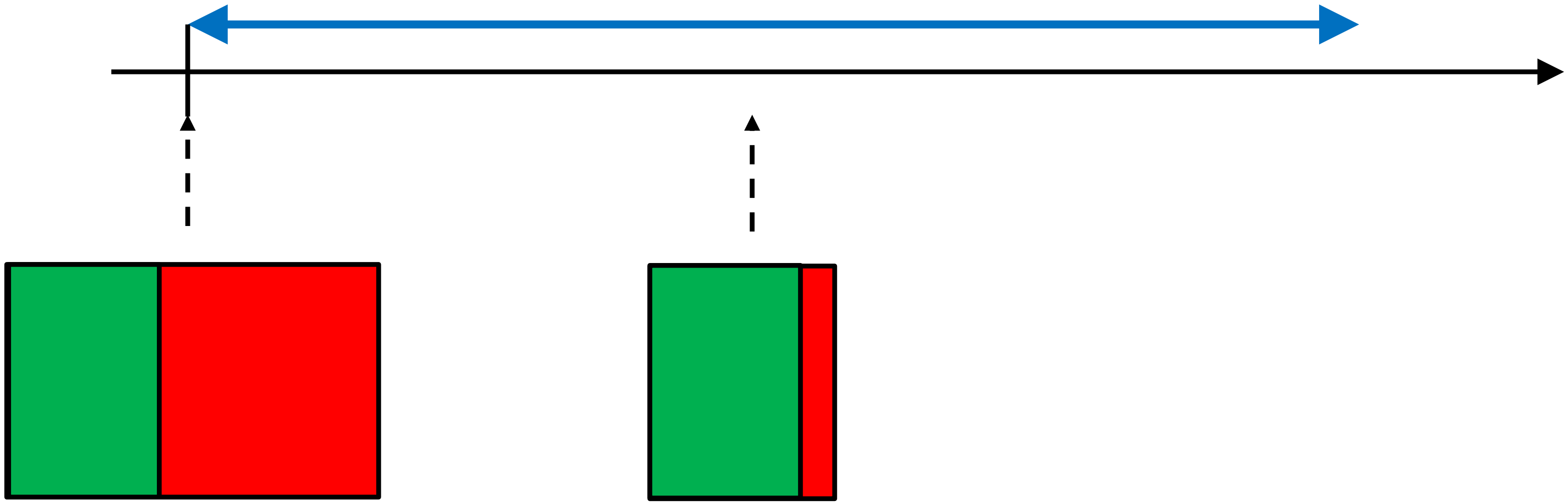
High Efficiency / Low Coverage



I have patched only vulns that matter

...but I have missed many vulns

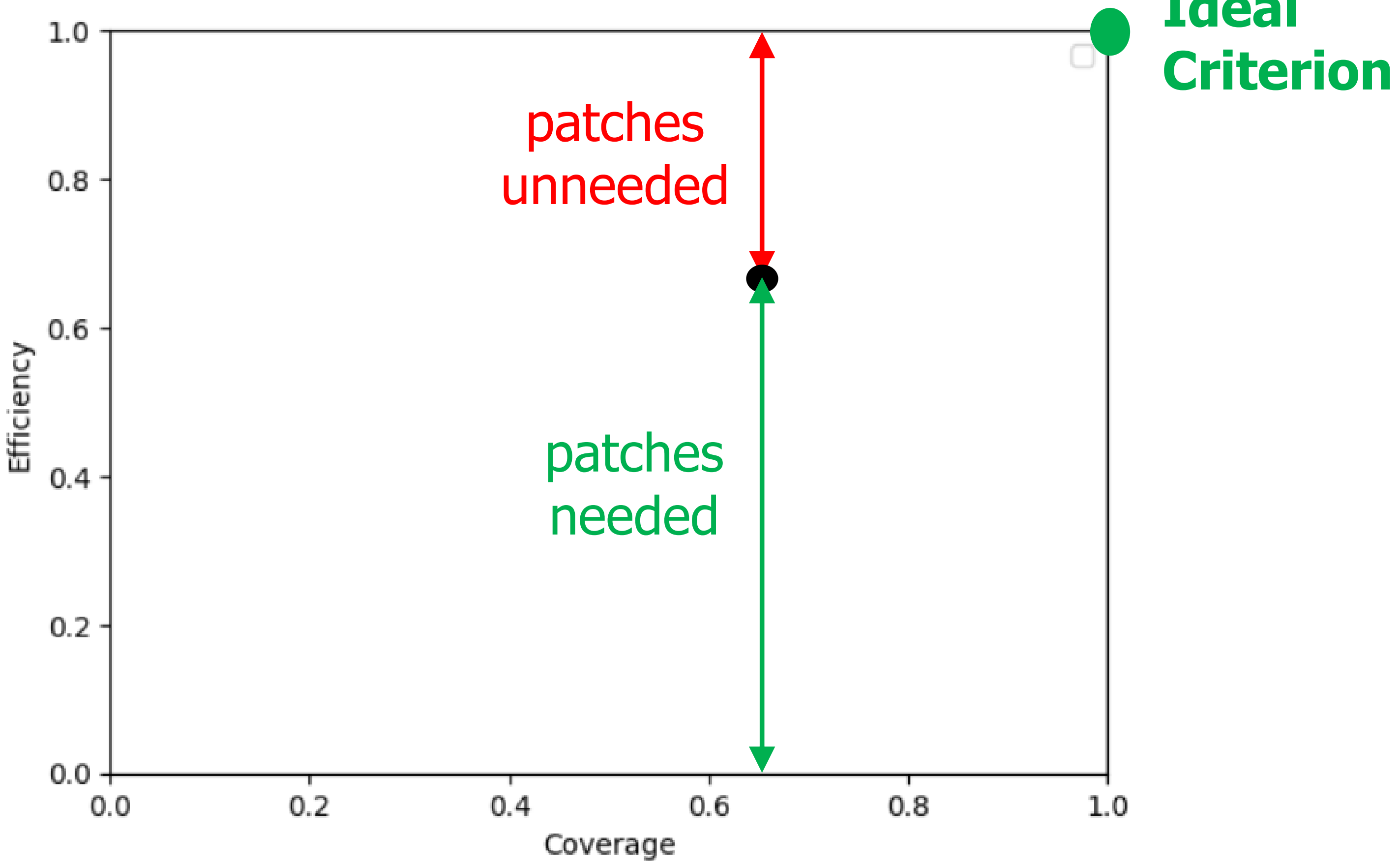
Low Efficiency / High Coverage



I have wasted lot of patching effort

...but I have covered nearly all vulns

Efficiency

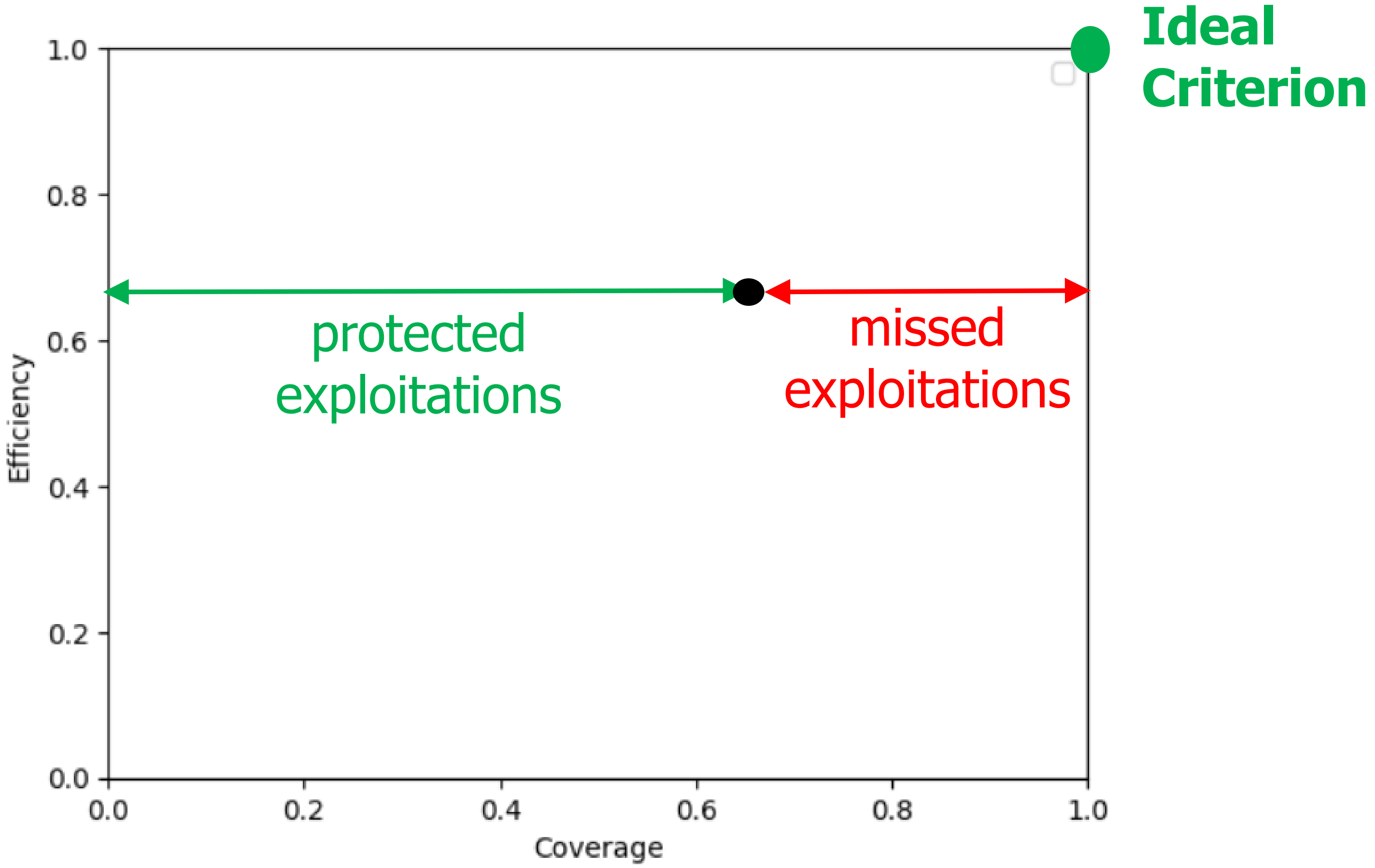


Ideal Criterion

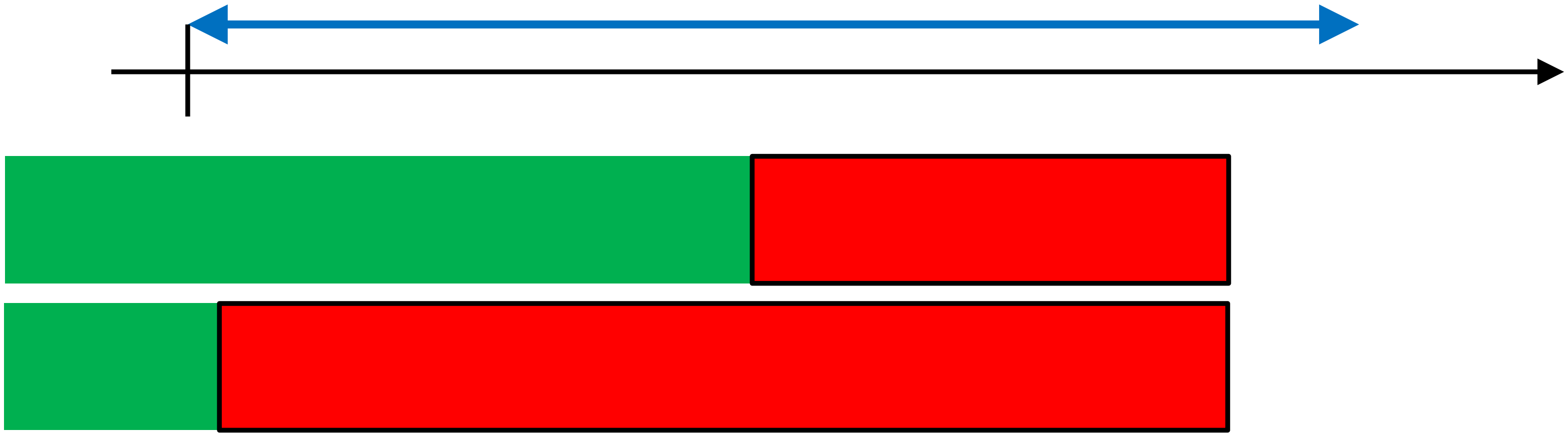
patches unneeded

patches needed

Coverage


1.0
0.8
0.6
0.4
0.2
0.0
Efficiency
0.0
0.2
0.4
0.6
0.8
1.0
Coverage
Ideal Criterion
protected exploitations
missed exploitations

Remark



- Coverage and Efficiency are **relative** indexes
- **Independent** of Patching Effort

- They all depend on the criterion for constructing the PatchingSet

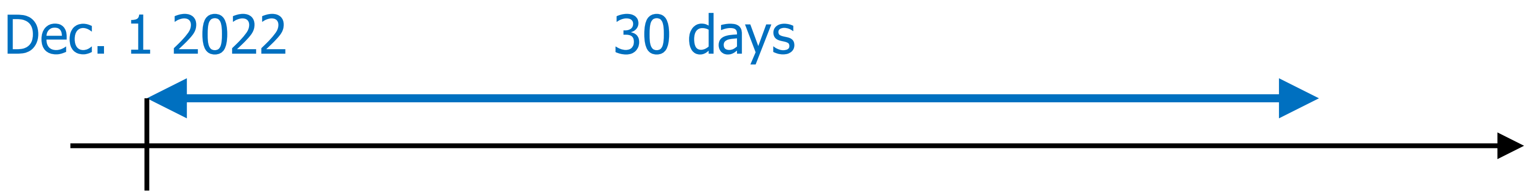
Problem Definition: Summary

- Criterion for **choosing which vulnerabilities to patch**
- Assessment indexes:
  - How good in defense (coverage)
  - How efficient (efficiency)
  - How costly (patching effort)

- **MANY** factors **not** assessed
- Given a certain vuln:
  - How **many systems** do I have with that vuln?
  - How **costly** is an **incident** based on that vuln?
  - How **likely** is that **I** will be attacked with that vuln?
  - ...

Exploit Prediction: Example Criteria

Experiment Scenario



Dec. 1 2022

30 days

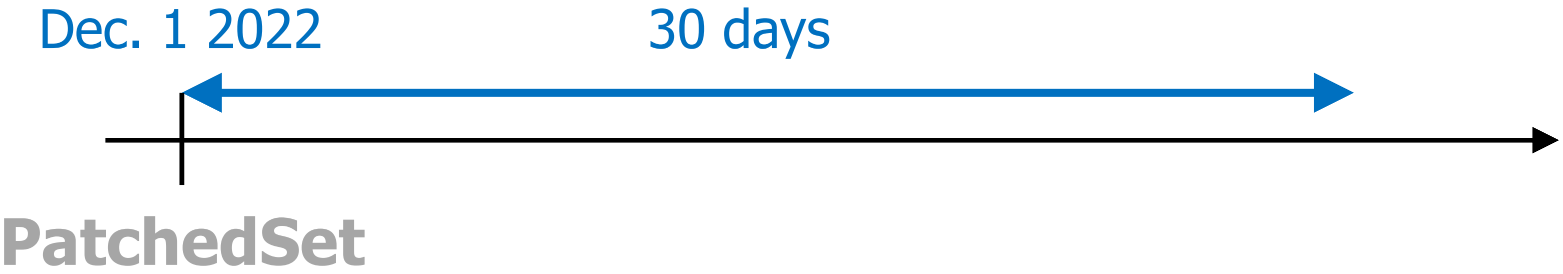
PatchedSet

ExploitedSet

Various Criteria

(≈8000 vulns)

Selection based on CVSS (I)



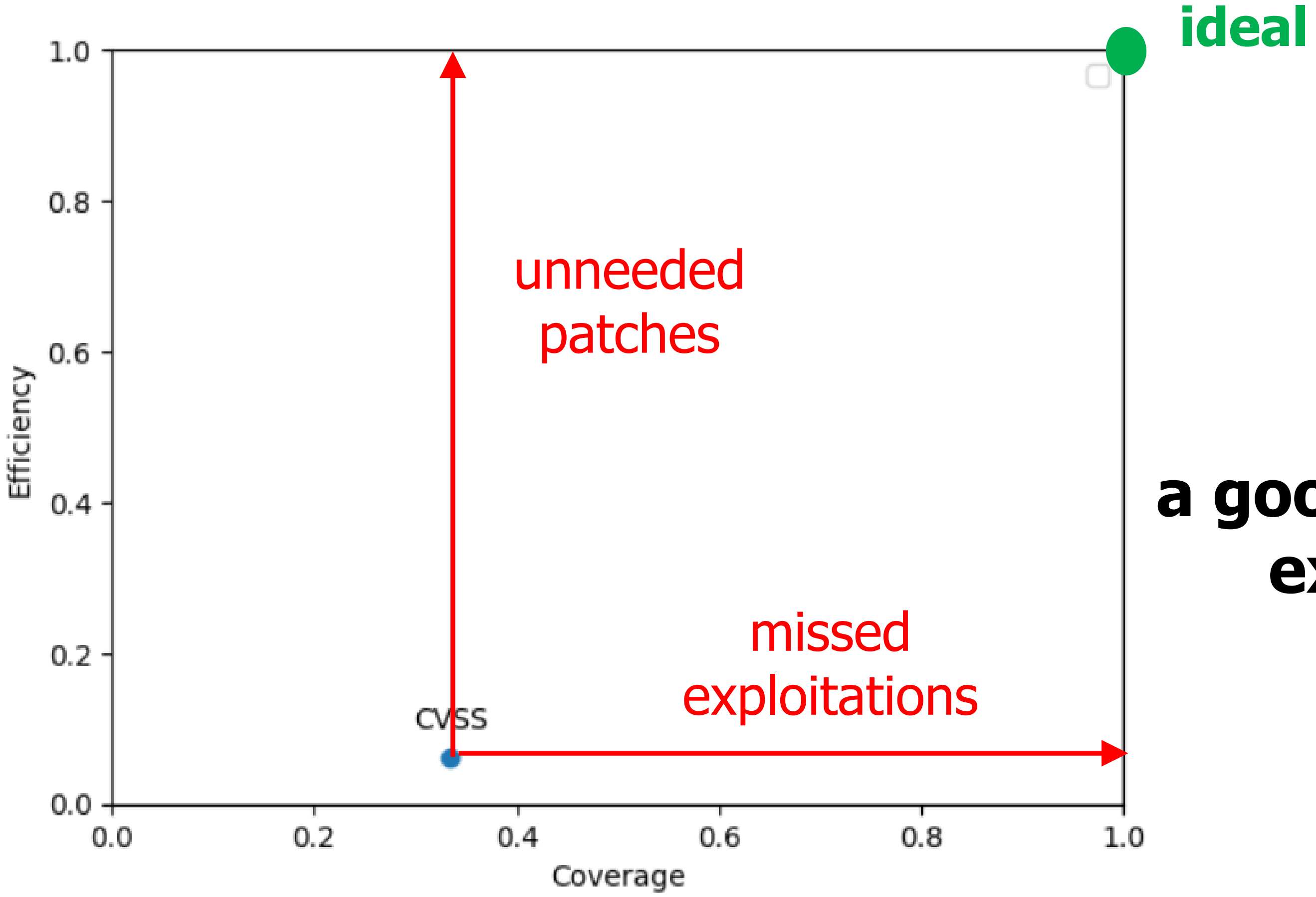
Dec. 1 2022

30 days

PatchedSet

- CVEs with **CVSS >= 9.1** (≈15% of all vulns)
- Patching Effort: ≈28000 vulns

Selection based on CVSS (II)



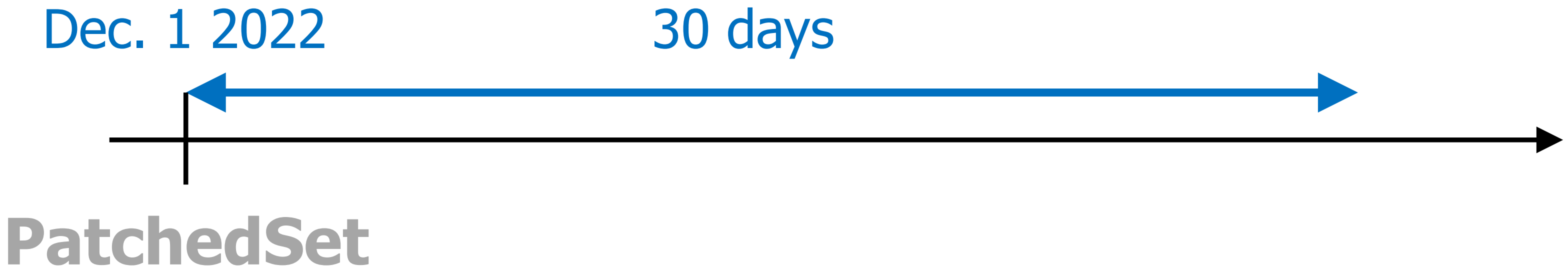
CVSS
is not
a good predictor of
exploitation

CISA - KEV

CYBERSECURITY
& INFRASTRUCTURE
SECURITY AGENCY

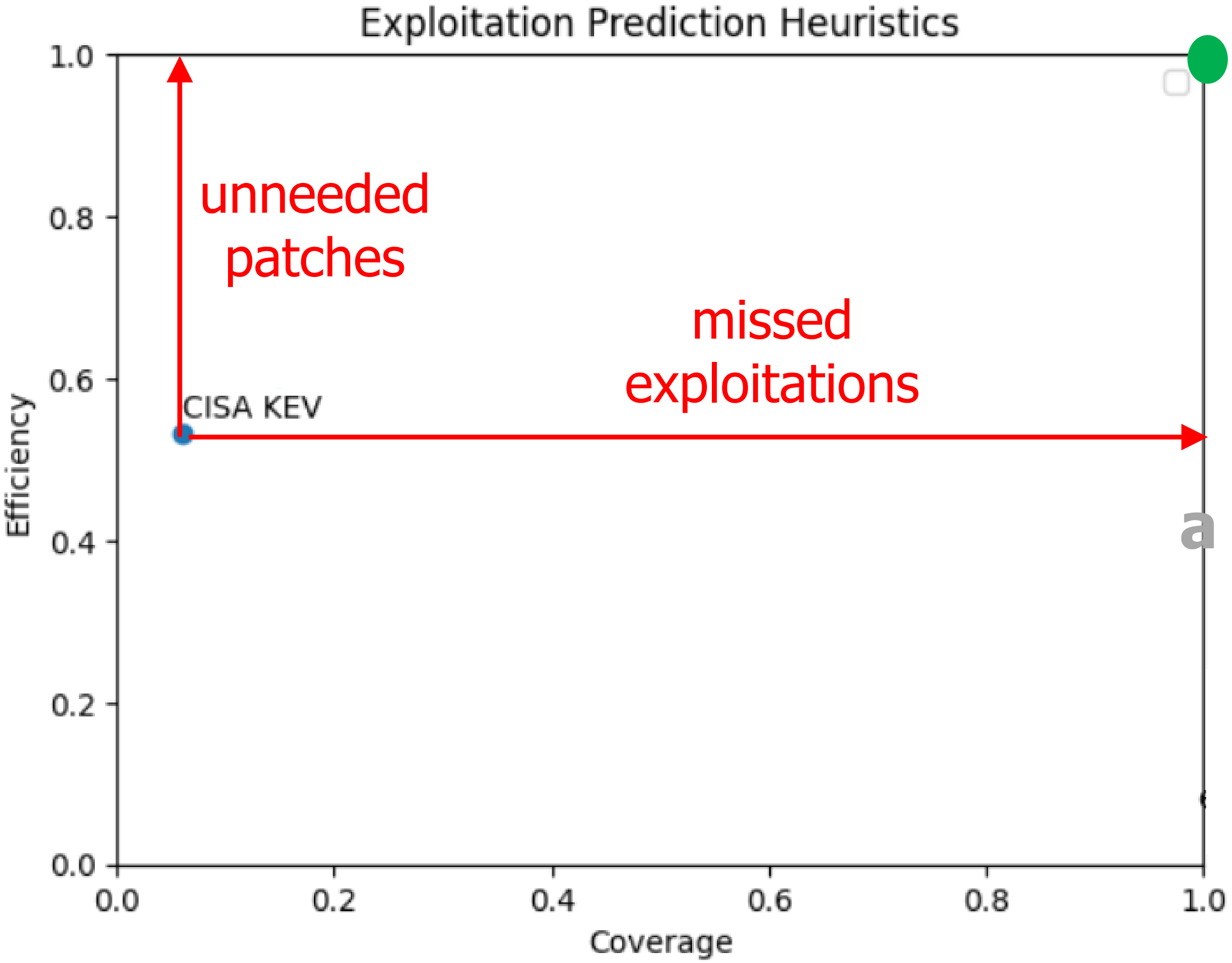
KNOWN EXPLOITED VULNERABILITIES CATALOG

Selection based on CISA-KEV (I)



- CVEs **in CISA-KEV** (≈0.5% of all vulns)
- Patching Effort: ≈900 vulns

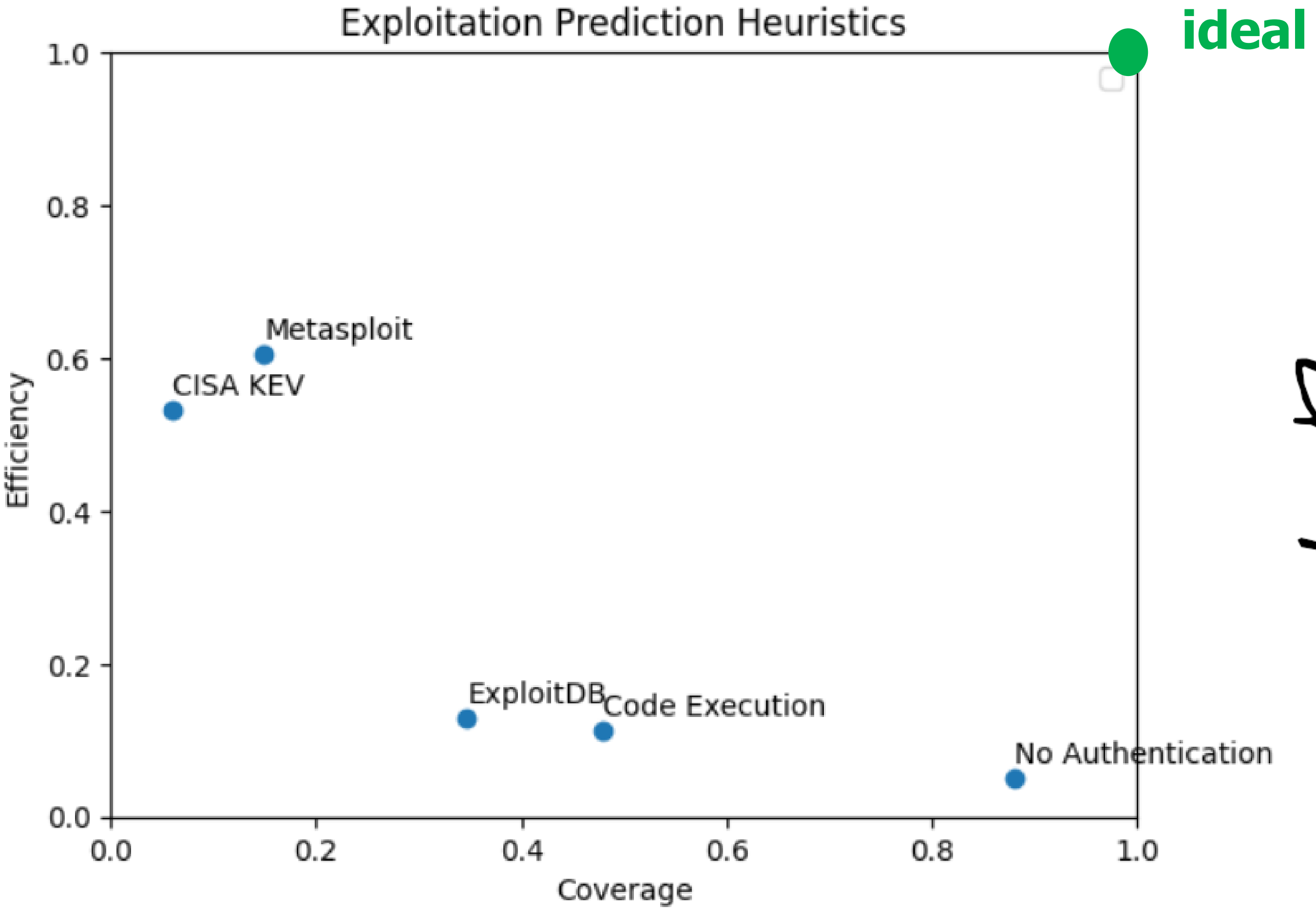
Selection based on CISA-KEV (II)



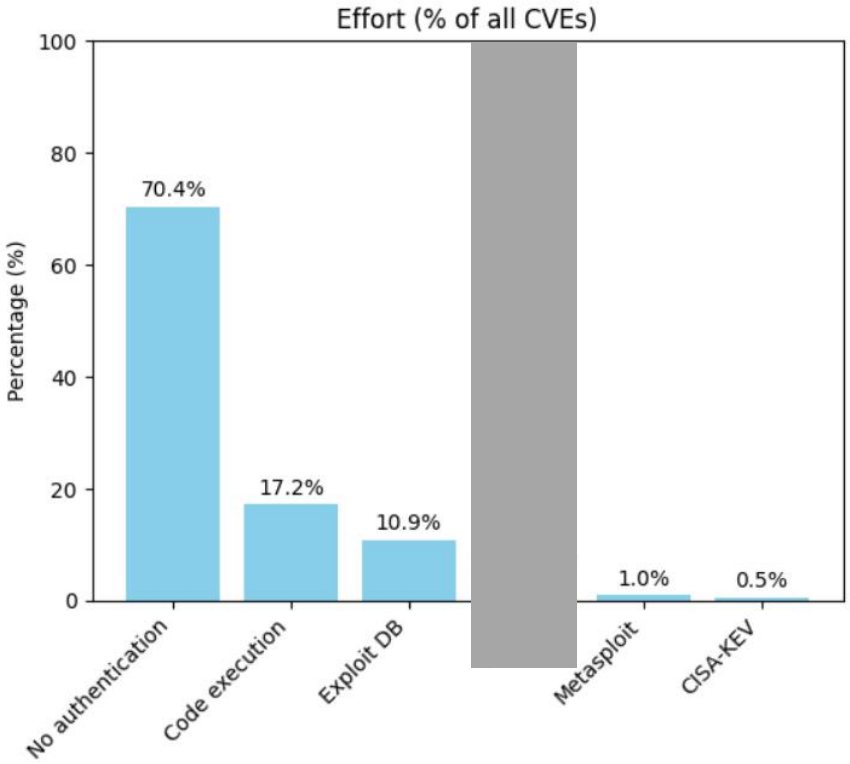
ideal

CISA-KEV
is not
a good predictor of
exploitation

Selection based on exploit properties (I)



Selection based on exploit properties (II)



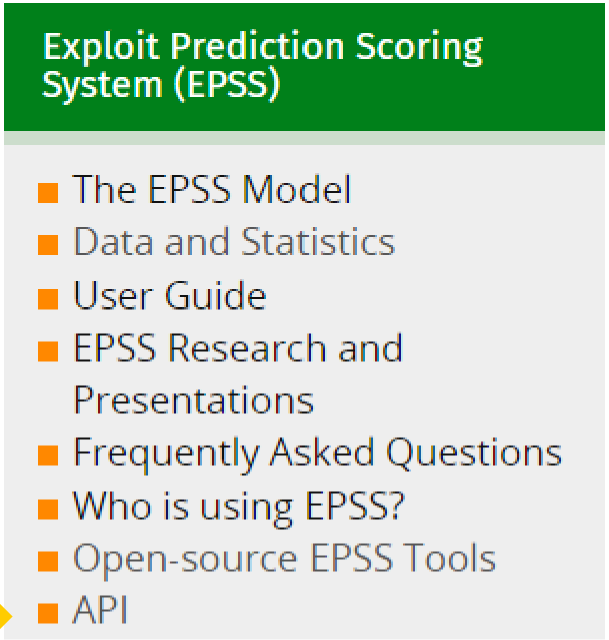
Exploit Prediction Scoring System (EPSS)

EPSS (I)

- EPSS(CVE-i, d):
  - Probability **estimate** that CVE-i will be exploited in [d, d+30]

- It changes **daily**
- Probability definition:
  #CVE-i exploitation **attempts worldwide** /
  #All CVE exploitation **attempts worldwide**

EPSS (II-a)

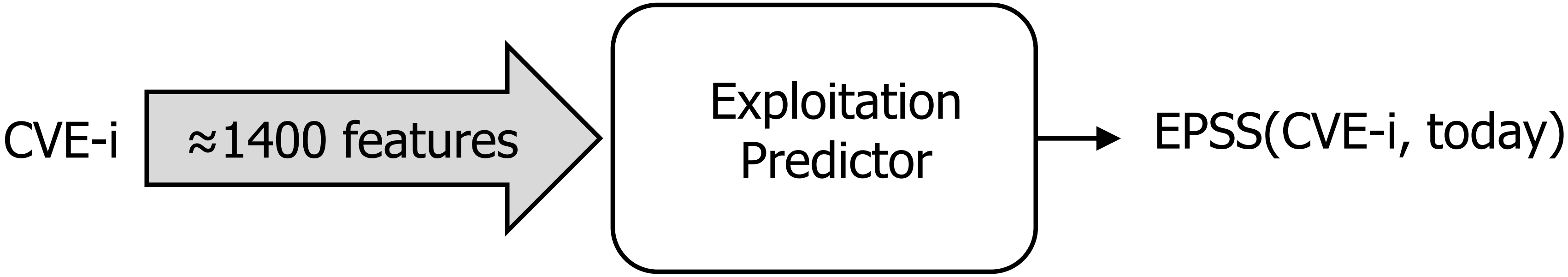


Exploit Prediction Scoring System (EPSS)

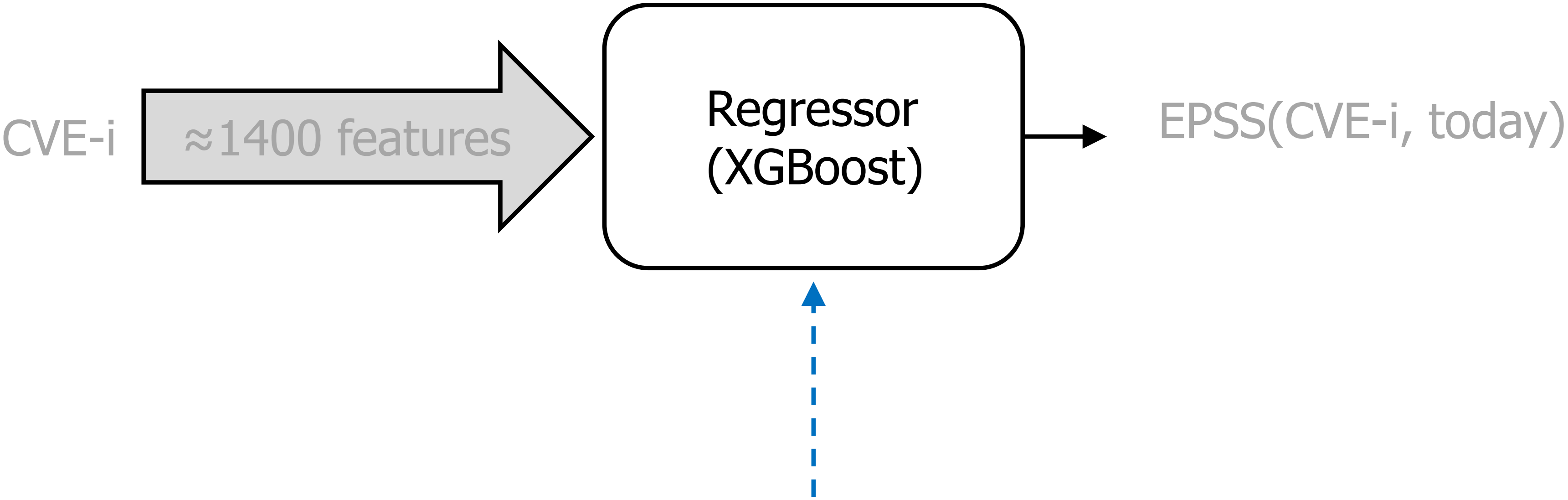
- The EPSS Model
- Data and Statistics
- User Guide
- EPSS Research and Presentations
- Frequently Asked Questions
- Who is using EPSS?
- Open-source EPSS Tools
- API

EPSS (II-b)

- Repeat **every day**:
  - For each CVE-i:
    - **Compute features** of CVE-i
    - **Estimate** its probability of exploitation in the next 30 days



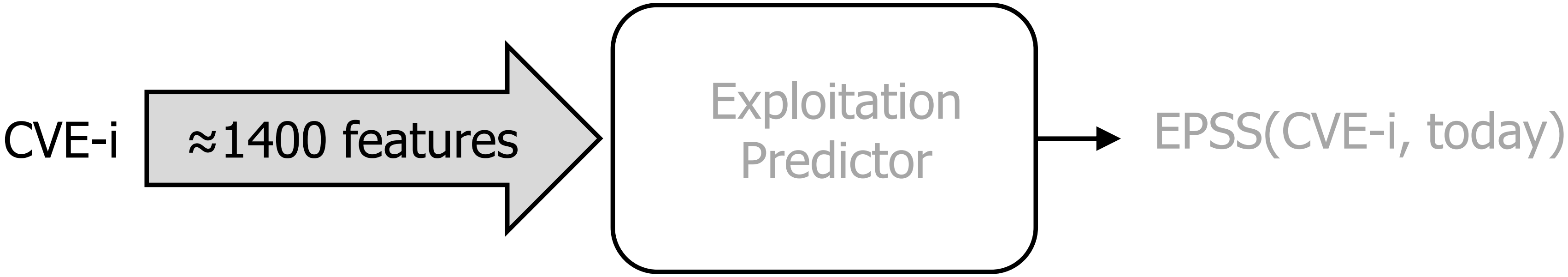
How does it work?



- Data driven model
- Trained on 1 year of observed exploitation activity

- March 2023: 3rd model refinement

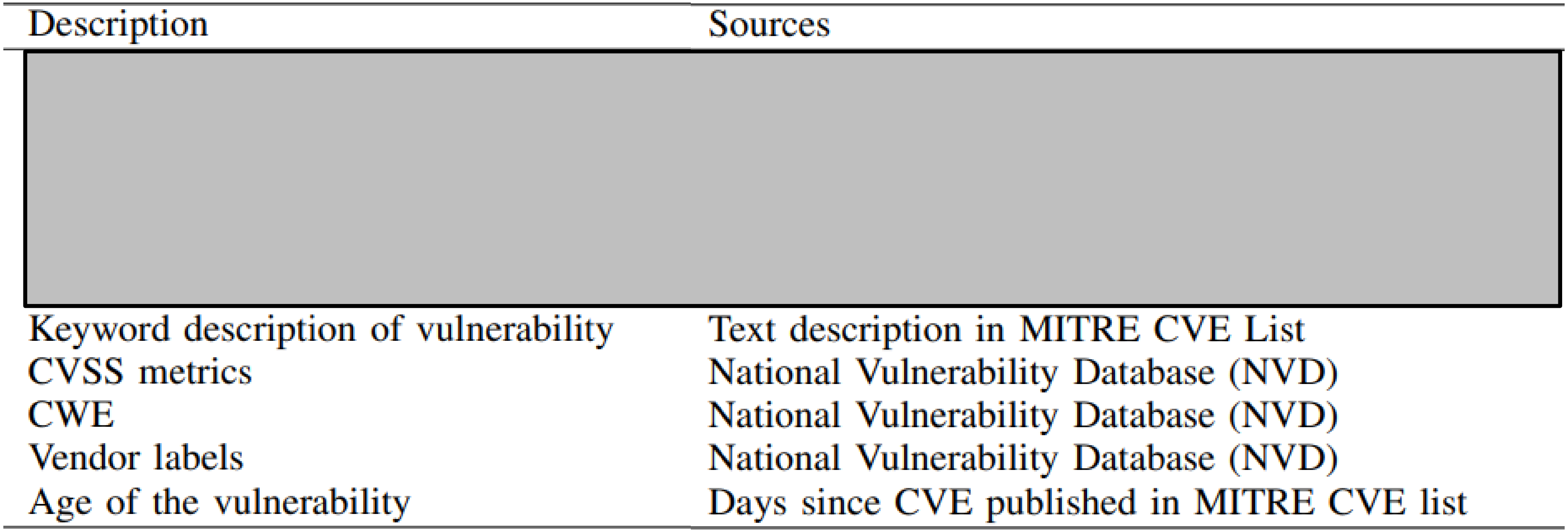
How is each vuln represented? (I)



- Array with 1400 elements
  - Numerical features
  - Categorical features (one-hot representation)

- Details out of scope
- Information sources in scope

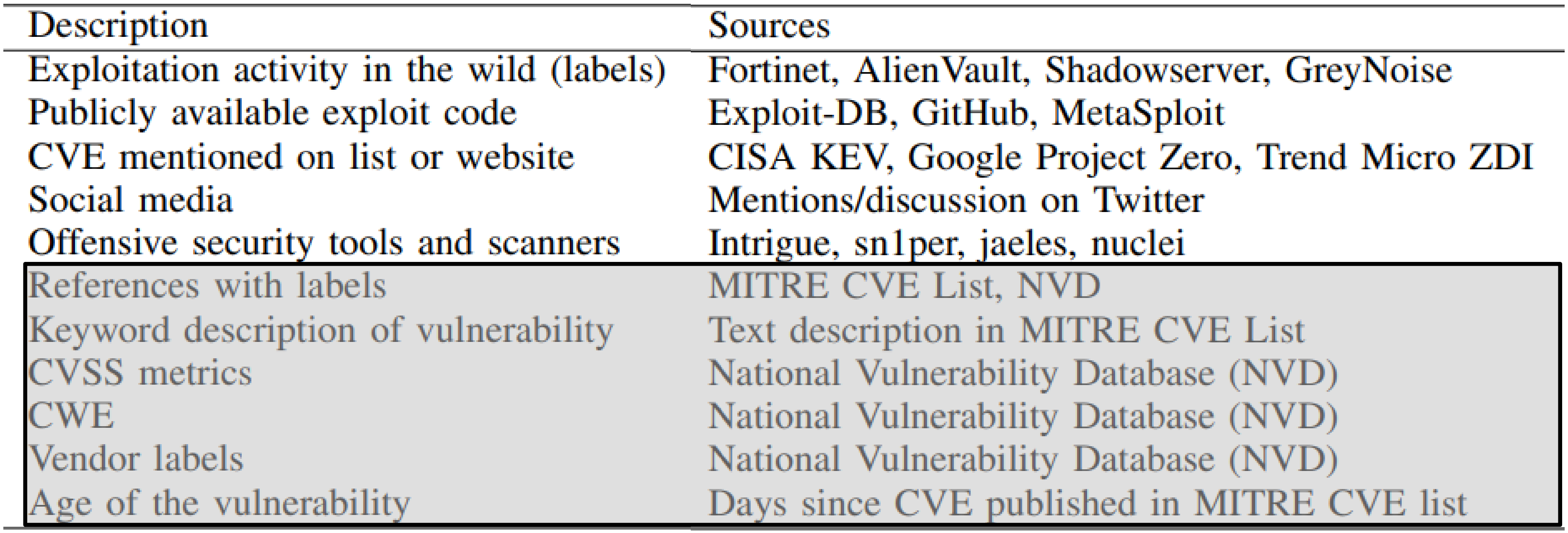
How is each vuln represented? (II)



Description	Sources
[redacted]	[redacted]
Keyword description of vulnerability	Text description in MITRE CVE List
CVSS metrics	National Vulnerability Database (NVD)
CWE	National Vulnerability Database (NVD)
Vendor labels	National Vulnerability Database (NVD)
Age of the vulnerability	Days since CVE published in MITRE CVE list

(more or less) intrinsic properties

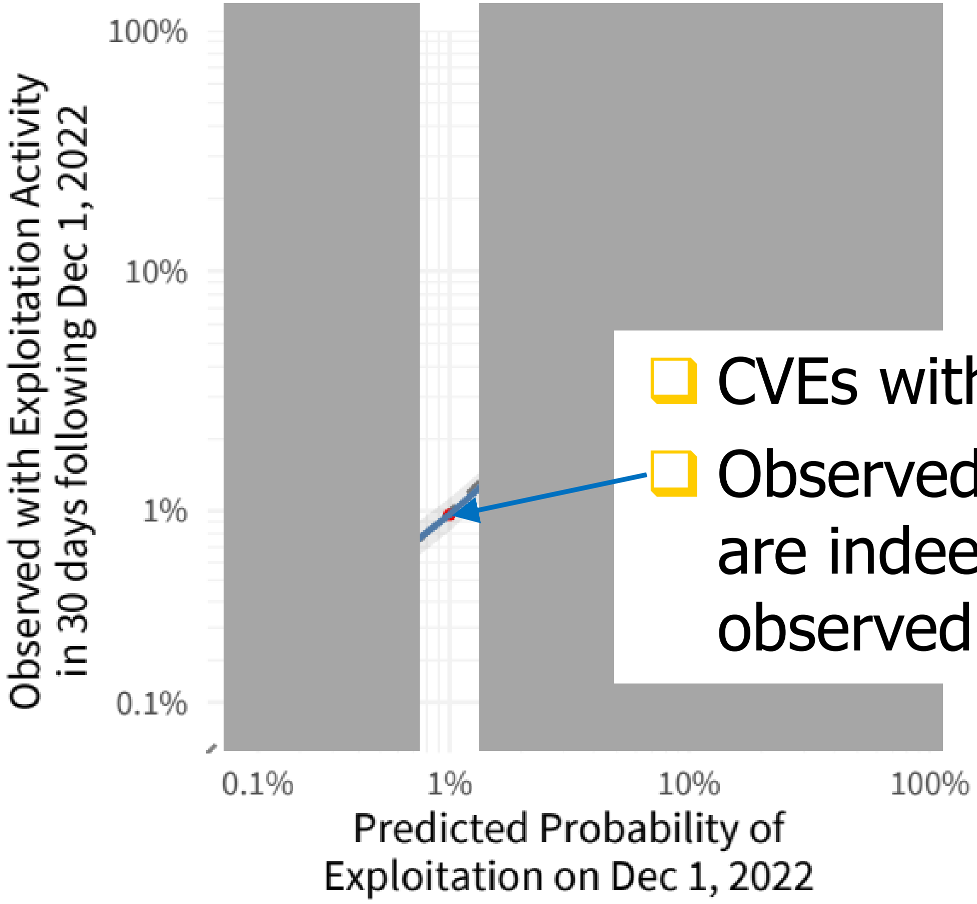
How is each vuln represented? (III)



Description	Sources
Exploitation activity in the wild (labels)	Fortinet, AlienVault, Shadowserver, GreyNoise
Publicly available exploit code	Exploit-DB, GitHub, MetaSploit
CVE mentioned on list or website	CISA KEV, Google Project Zero, Trend Micro ZDI
Social media	Mentions/discussion on Twitter
Offensive security tools and scanners	Intrigue, sn1per, jaeles, nuclei
References with labels	MITRE CVE List, NVD
Keyword description of vulnerability	Text description in MITRE CVE List
CVSS metrics	National Vulnerability Database (NVD)
CWE	National Vulnerability Database (NVD)
Vendor labels	National Vulnerability Database (NVD)
Age of the vulnerability	Days since CVE published in MITRE CVE list

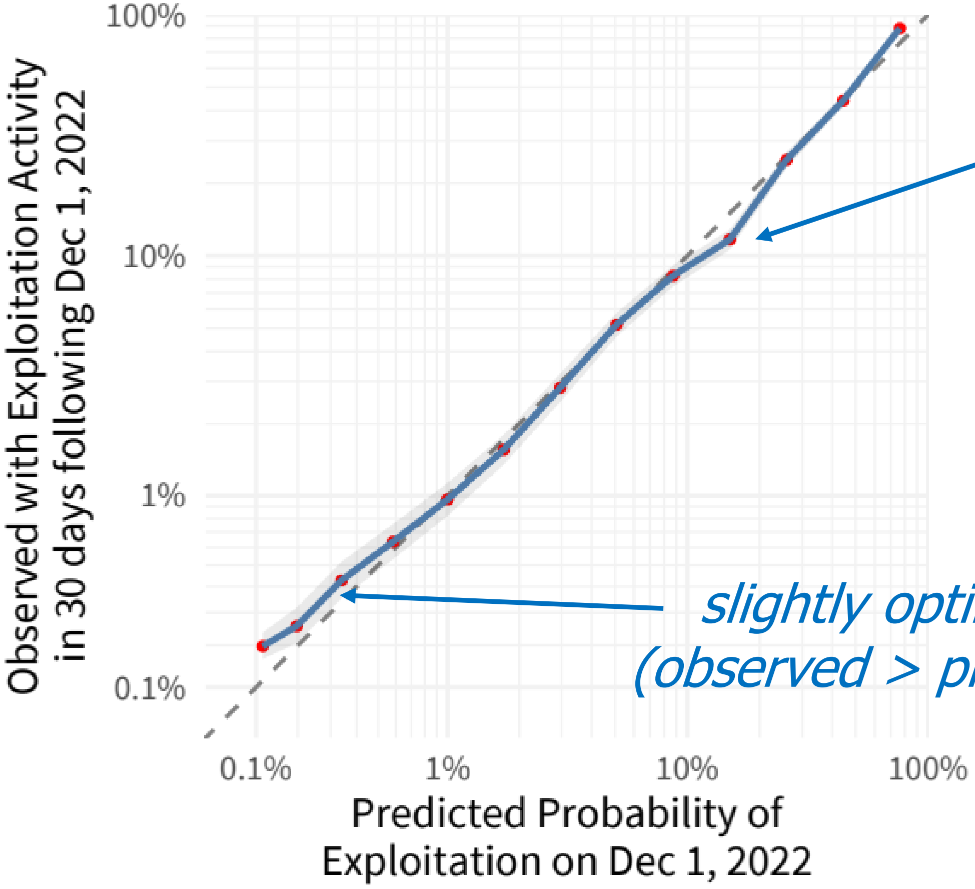
- Summary of "**what people are saying** of this vuln"
- Updated **daily**

P_observed(CVE-i) vs P_predicted(CVE-i)



- ❑ CVEs with EPSS = 0.01
- ❑ Observed exploitation attempts are indeed 0.01 of all the observed exploitation attempts

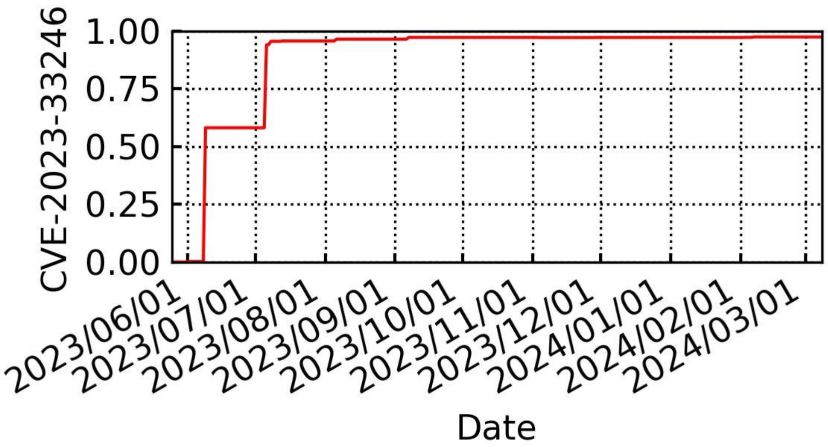
P_observed(CVE-i) ≈ P_predicted(CVE-i)



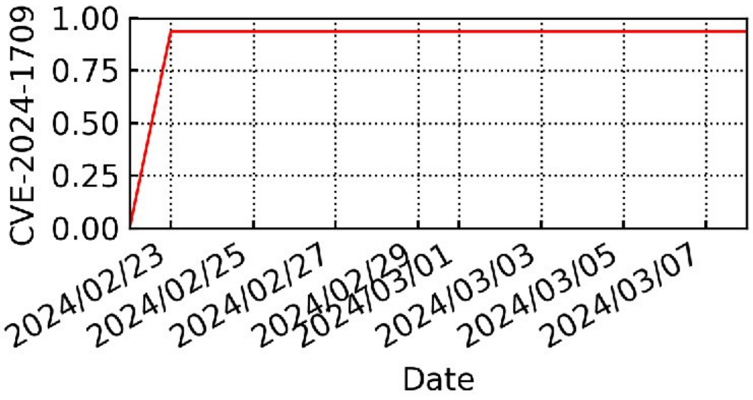
slightly pessimistic

slightly optimistic
(observed > predicted)

EPSS evolution: Examples (I)

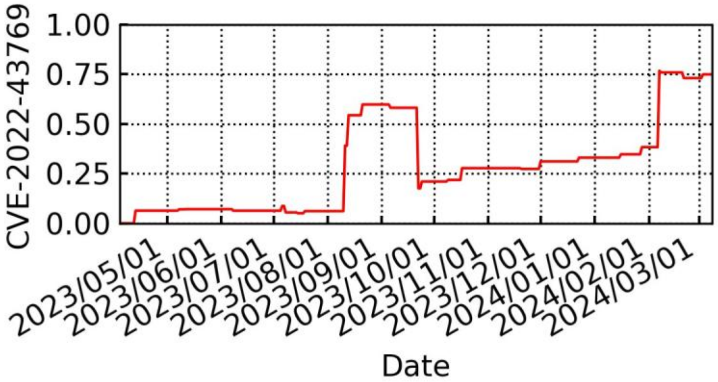


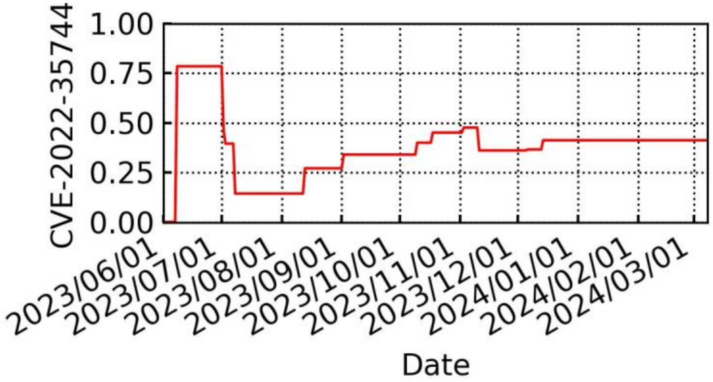
- Significant growth **after 1 day**
- ...and then again on the next day
- Heavily exploited for more than 6 months



- **Immediately** exploited heavily, for several weeks

EPSS evolution: Examples (II)





- Temporal evolution may often be:
  - Very "irregular"
  - Very hard to predict (even in the short term)

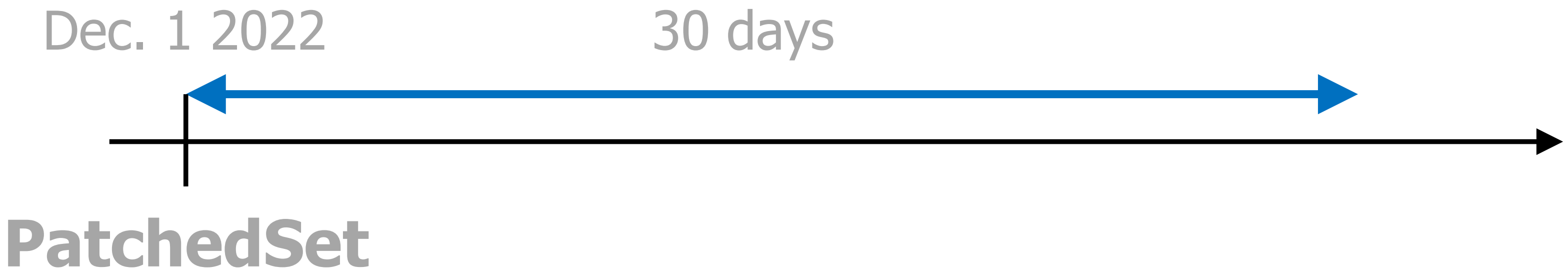
Remark

- E**P**SS(CVE-i, d):
  - Probability **estimate** that CVE-i **will be** exploited in [d, d+30]
- It is called a **predictor**

- Summary of "**what people are saying** of this vuln"
- Updated **daily**
- It actually acts **retrospectively**
- ...and it may have **delays of several days**

EPSS for Exploit Prediction: Coverage and Efficiency?

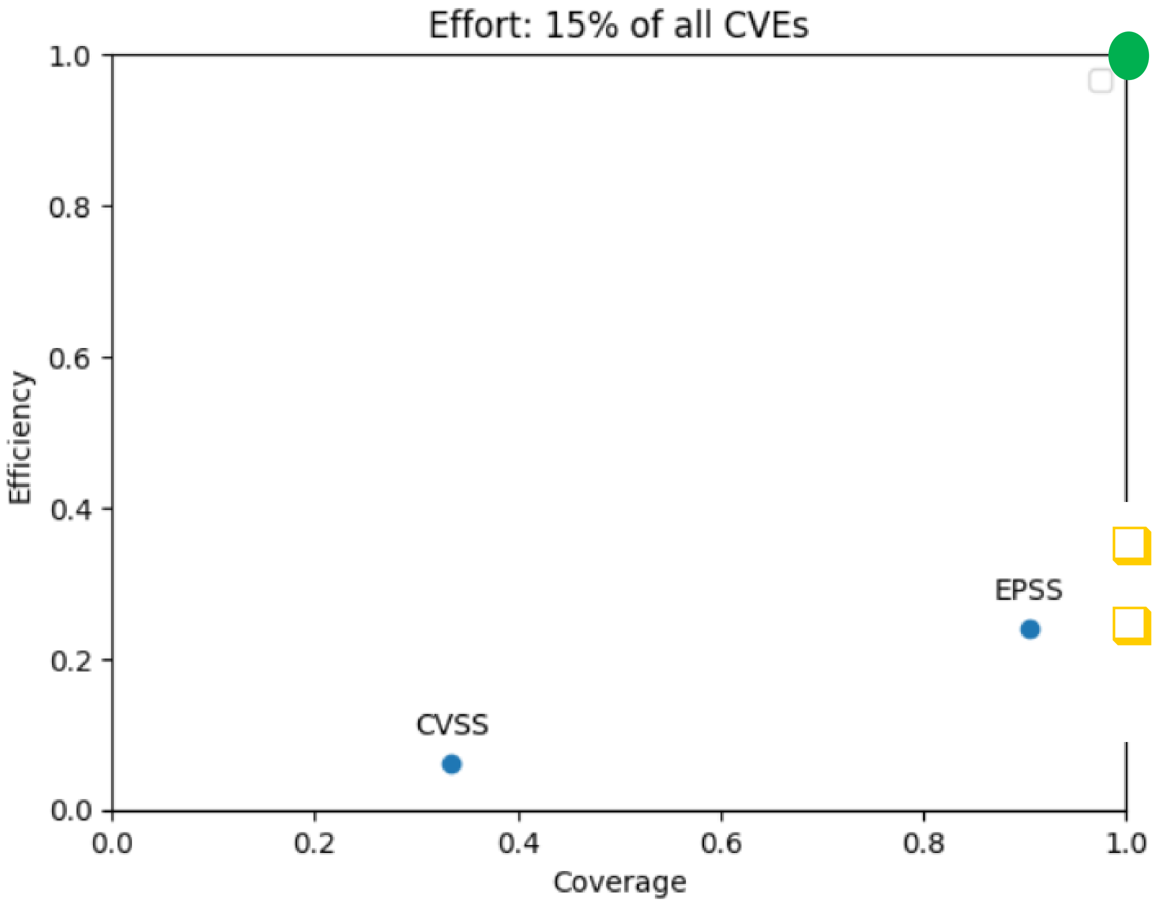
EPSS vs CVSS: Same Patching Effort (I)



- Set1: CVSS(CVE-i) >= 9.1 (15% of all CVEs)
- Set2: EPSS(CVE-i) >= **0.022** (15% of all CVEs)

- **Identical Patching Effort**
- Efficiency?
- Coverage?

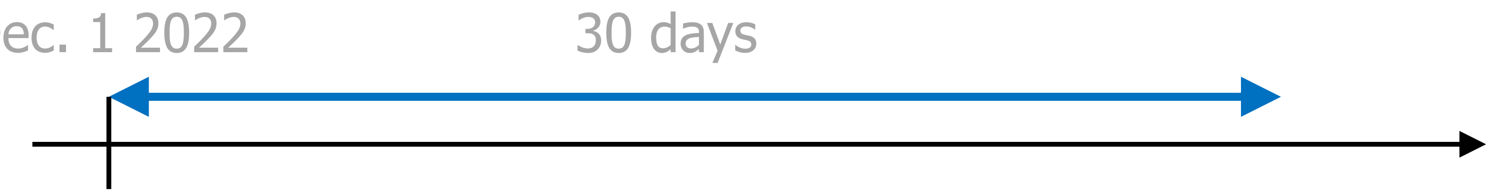
EPSS vs CVSS: Same Patching Effort (II)



ideal

- ❑ Excellent coverage
- ❑ More efficient than CVSS (But not very efficient)

EPSS vs CVSS: Same Coverage (I)

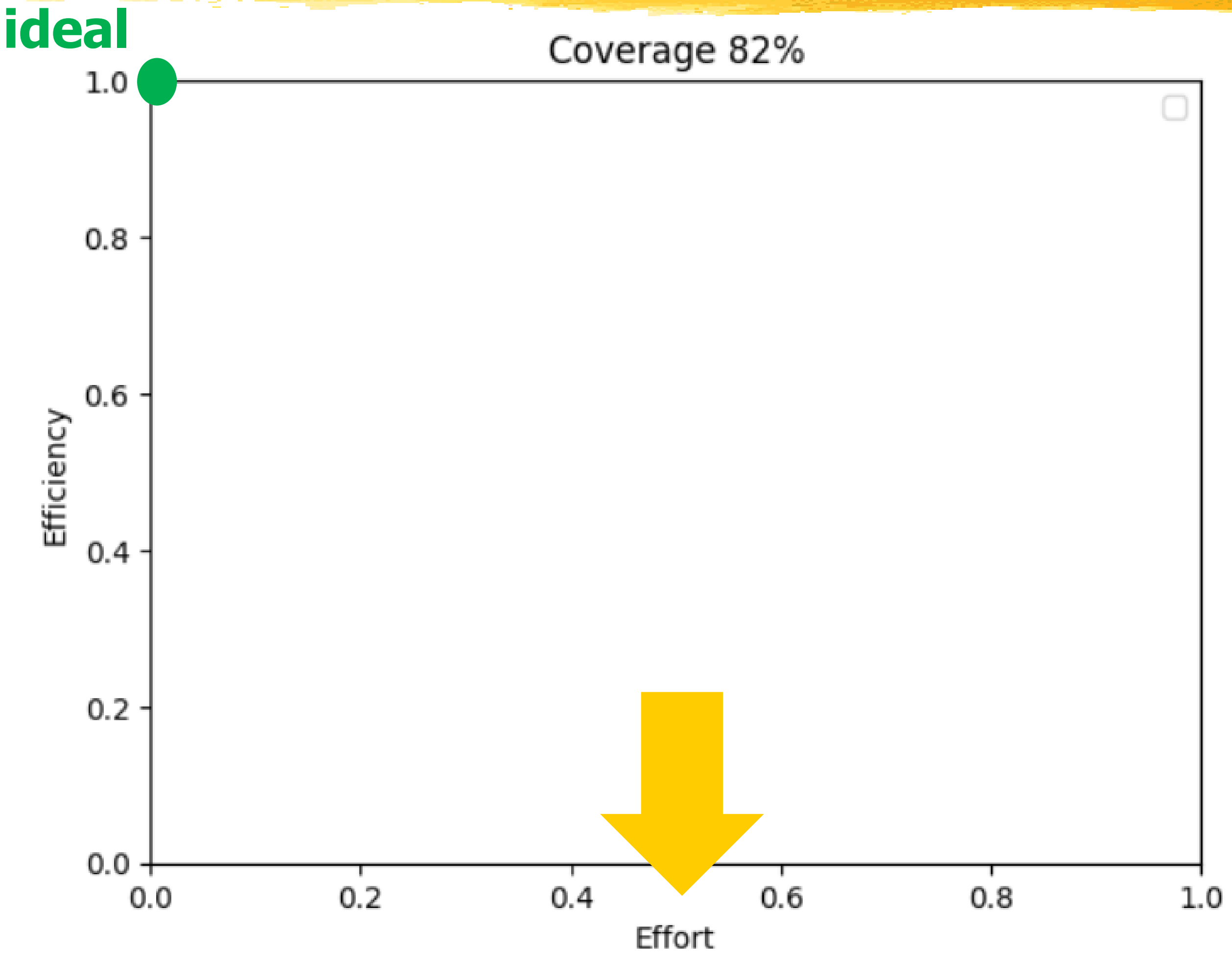


PatchedSet

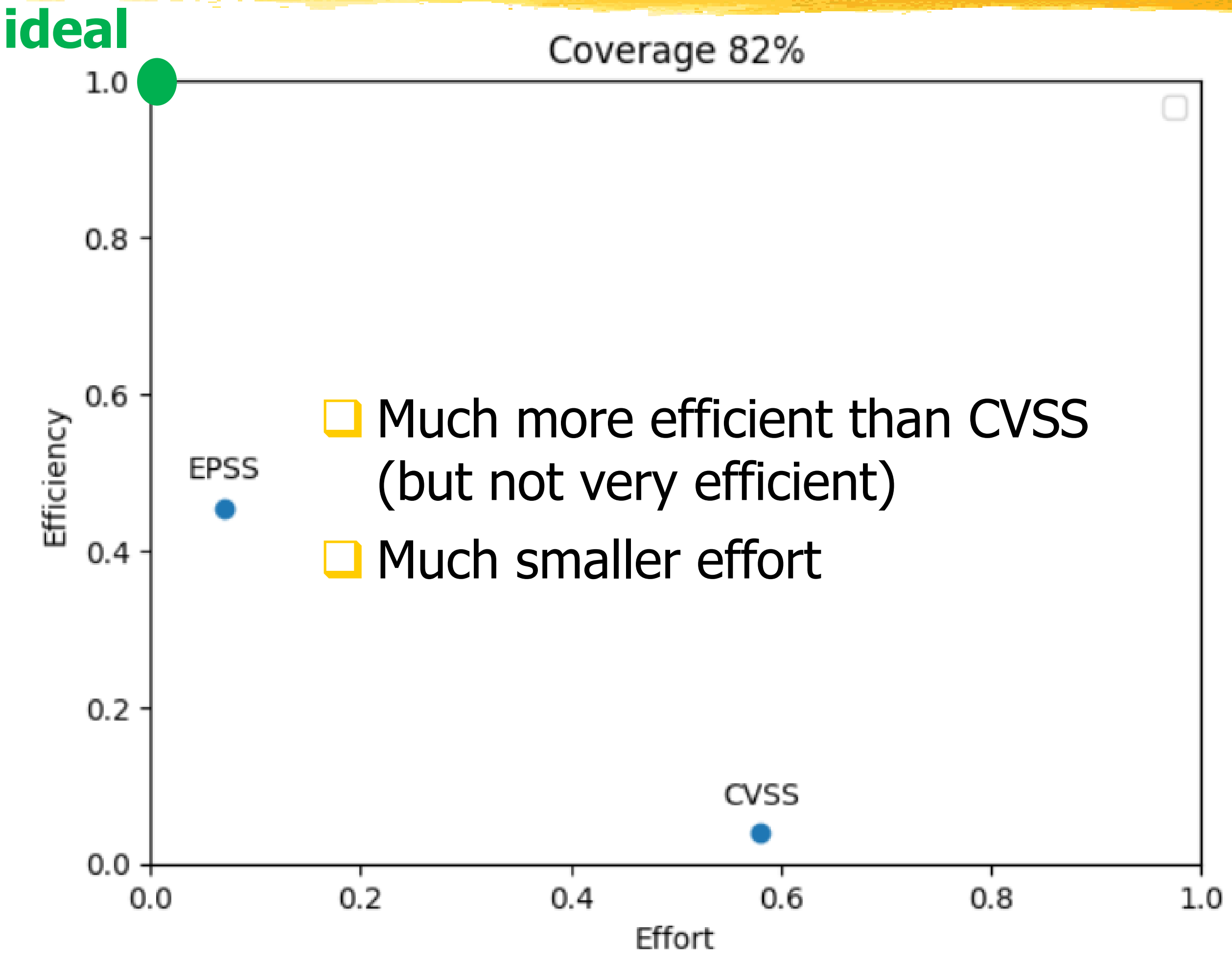
- Set1: CVSS(CVE-i) >= 7 (Coverage 82%)
- Set2: EPSS(CVE-i) >= **0.088** (Coverage 82%)

- **Identical Coverage**
- Efficiency?
- Patching Effort?

EPSS vs CVSS: Same Coverage (II-a)

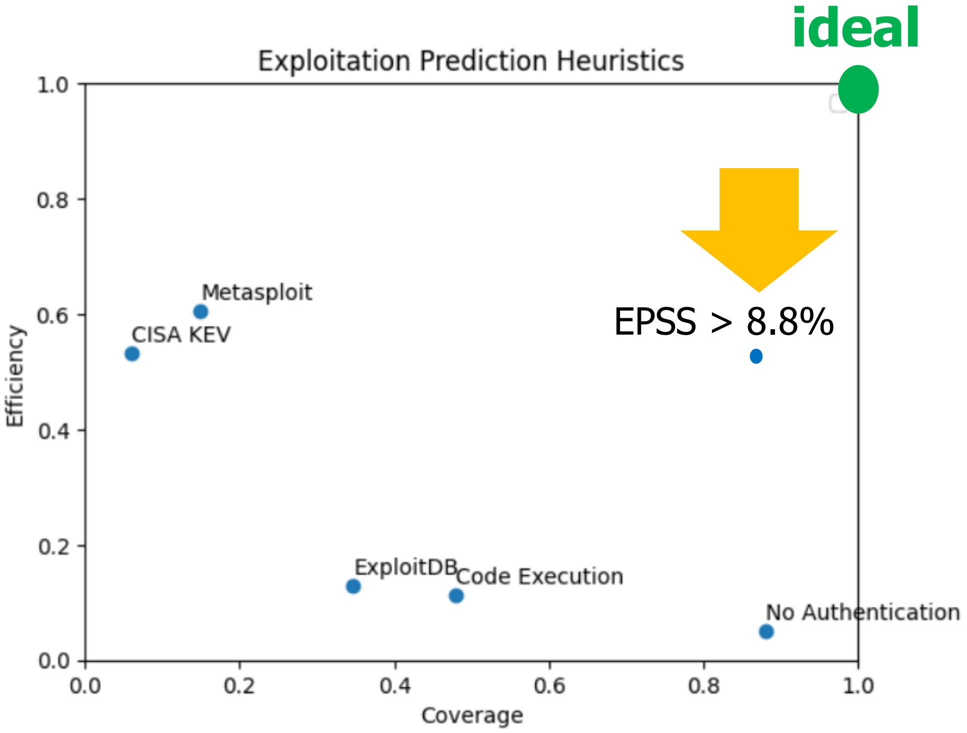


EPSS vs CVSS: Same Coverage (II-b)

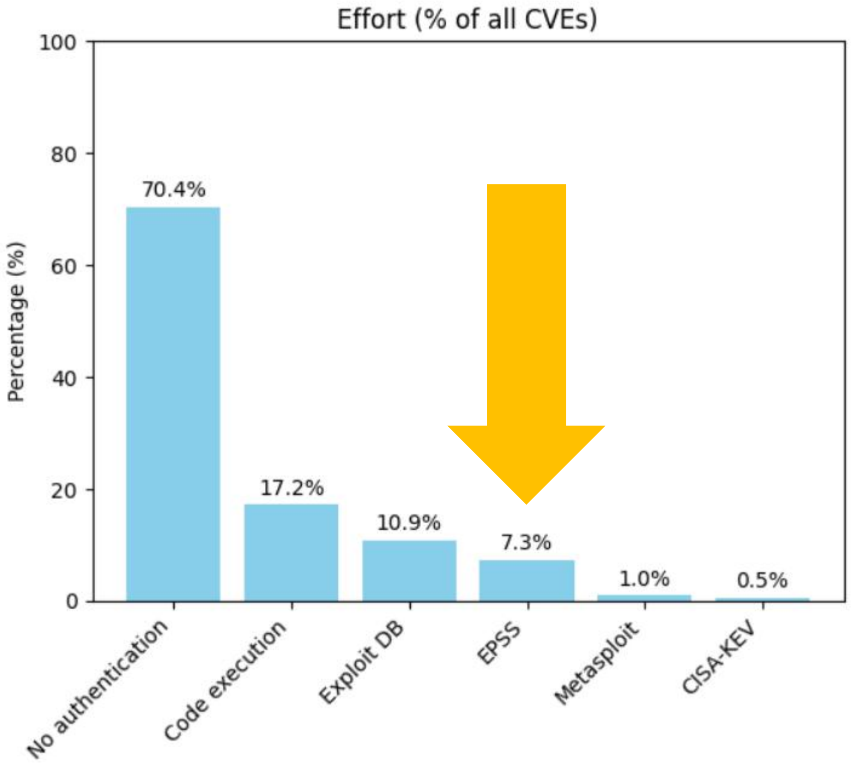


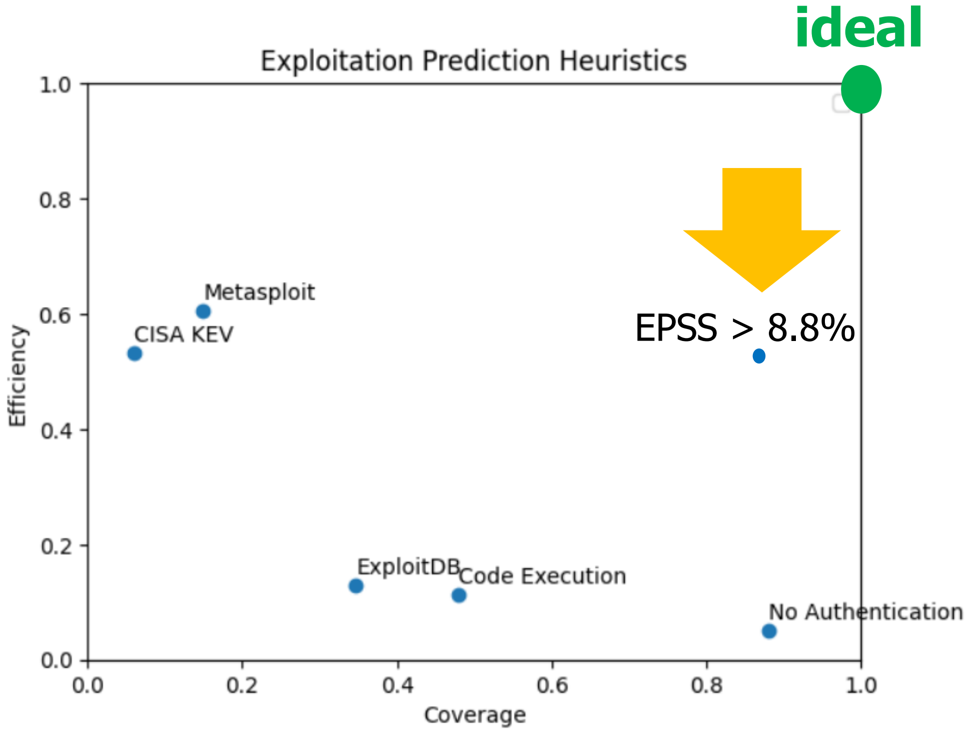
- Much more efficient than CVSS (but not very efficient)
- Much smaller effort

EPSS vs Heuristics (I)



EPSS vs Heuristics (II)




ideal

EPSS > 8.8%