UNIVERSITY OF TRENTO

SECONOMICS

Exploitation in the wild: what do attackers do, and what should(n't) we care about.

Luca Allodi, Fabio Massacci
University of Trento, Italy.
$name.$surname@unitn.it

Outline

- Introduction **(3 slides)**
  - Vulnerability Management guidelines: CVSS
  - What do the IT Sec Managers need: research question
- Vulnerability landscapes **(5 slides)**
  - The good guys
  - Most bad guys
  - Our baseline: data
  - Reality on attacks, according to the data
- Observational analysis of CVSS scores **(5 slides)**
  - CVSS distributions
  - Map of vulnerabilities, exploits and CVSS scores: CVSS not good
- What makes the CVSS so inaccurate? **(15 slides)**
  - Inspection of CVSS subscore distributions
  - Case controlled study: CVSS as a test for exploitation
  - Relative diminishment in risk with vulnerability patching
- Conclusions

Introduction

Vulnerabilities guidelines

- US Government SCAP Protocol for vulnerability remediation [Scarfone 2010]

> *"Organizations should use CVSS base scores to assist in prioritizing the remediation of known security-related software flaws based on the relative severity of the flaws."*

Vulnerabilities guidelines

- US Government SCAP Protocol for vulnerability remediation [Scarfone 2010]

> *"Organizations should use CVSS base scores to assist in prioritizing the remediation of known security-related software flaws based on the relative severity of the flaws."*

➡ bother with every software vulnerability, use CVSS to prioritize your work

Don't cite me on that (they said)

- *"My job is **the** professional nightmare: if everything goes well, I am not doing anything. If something goes badly wrong, I am fired."* – Security Manager of big Italian player in sw industry
- *"Just acknowledging there is a bug costs hundreds of euros"* – Representative of EU leader in sw management
- *"You are crazy if you think I'll install all the patches"* – IT Admin of big US telecommunication company

Vulnerabilities: research question

- What the CIO would like to know
  - If I follow SCAP or equivalent guidelines, how much will my final risk decrease?

- A clear value proposition:
  - if we fix high CVSS vulns we decrease risk by +43%
  - if we fix all medium CVSS only raises to +48%
    - → +5% more is not worth the extra money, maybe even +43% is not worth

Vulnerabilities: landscapes

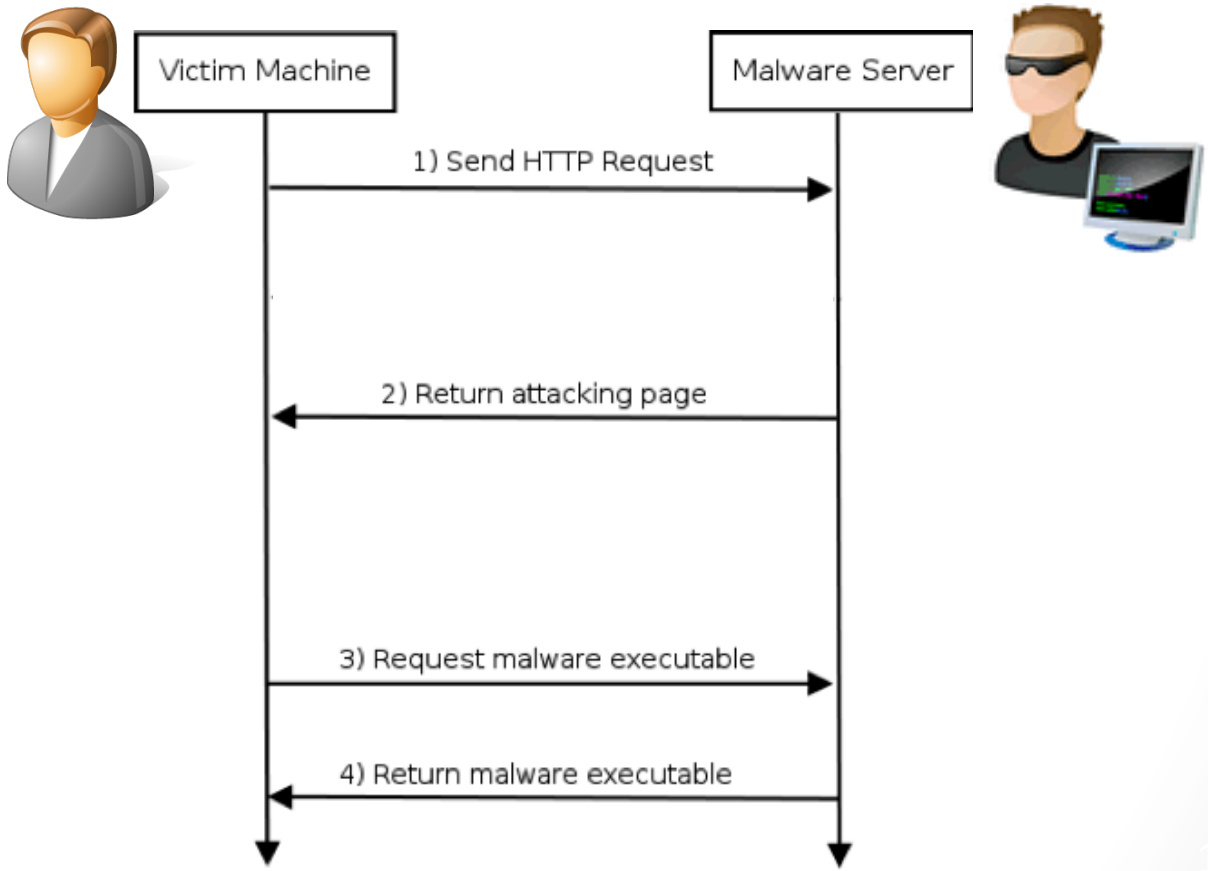
Vulnerabilities: the good guys

- Databases for vulnerabilities:
  - Lots of Vulnerabilities are published daily
  - NVD runs at 50K
  - CVSS scoring system is now drafting V.3
- Databases for exploits:
  - Vendors' "Bounty programs"
  - iDefender, TippingPoint acquisition program
  - "Responsible Disclosure" debate
- Analysis of complete protection against a powerful adversary
  - Classic model of the attacker [Dolev, Schneier…]

➢ Fix all vulnerabilities or die

Vulnerabilities: most bad guys

- Automated web attacks represent 2/3 of final threat for users [Google 2011],[Grier 2012]



Vulnerabilities: most bad guys

- Automated web attacks represent 2/3 of final threat for users [Google 2011],[Grier 2012]

Средний пробив на связке: **10-25%**
* Пробив указывается приблизительный, может отличаться и зависит напрямую от вида и качества траффика.

* Отстук стандартный, даже чуть выше стандартного:
> Зевс = 50-60%
> Лоадер = 80-90%

Exploitation success rate
*Rate highly depends on traffic quality

Цена последней версии 1.6.x:
> Стоимость самой связки = 2000$
> Чистки от АВ = от 50$
> Ребилд на другой домен/ИП = 50$
> Апдейты = от 100$
* Связка с привязкой к домену или IP .

Latest prices

Additional services

Связь:
> ICQ: **9000001**
> Jabber: Exmanoize@xmpp.jp

Рабочий график:
> понедельник - суббота
> с 7 до 17 по мск.

Vendor's contacts
Working hours:
- Monday-Saturday
- 7am to 5pm (Moscow time)

CVSS score

23.03.2011, 19:44

Апдейт до версии "***Eleonore Exp v1.6.5***"

В состав связки входят следующие эксплойты:
> CVE-2006-0003 (MDAC)
> CVE-2006-4704 (WMI Object Broke)
> CVE-2008-2463 (Snapshot)
> CVE-2010-0806 (IEpeers)
> CVE-2010-1885 (HCP)
> CVE-2010-0188 (PDF libtiff mod v1.0)
> CVE-2011-0558 (Flash <10.2)
> CVE-2011-0611 (Flash <10.2.159)
> CVE-2010-0886 (Java Invoke)
> CVE-2010-4452 (Java trust)
*Виста и 7ка бьется

Vulnerabilities: our baseline

- **NVD**
  - The universe of vulnerabilities
- **EXPLOIT-DB**
  - Exploits published by security researchers
- **EKITS** (The black markets)
  - 1.5 years of study of the black markets
  - Automated monitoring of exploit kits and new CVEs
  - 90+ exploit kits from the black markets
- **SYM**
  - Vulnerabilities actually exploited in the wild
  - Browser/Plugins 14% – Server 22% – App. 24%
  - Solaris, MacOs, Linux and others are included

dataset	volume
NVD	49.624
EDB	8.189
EKITS	126
SYM	1.289

Reality so far

- The "Classic" Attacker Model looks wrong
  - Few exploited vulnerabilities
  - Big chunk of risk from a bunch of vulnerabilities
  - ~~Fix all vulnerabilities or die~~ → waste of money?

- But CIO can't wait:
  - Use a Security Configuration Management Product!
  - 30+ products: Microsoft, Dell, HP, VMWare, McAfee, Symantec etc..
  - Based on CVSS (Common Vuln. Scoring System)

Observational analysis of CVSS scores

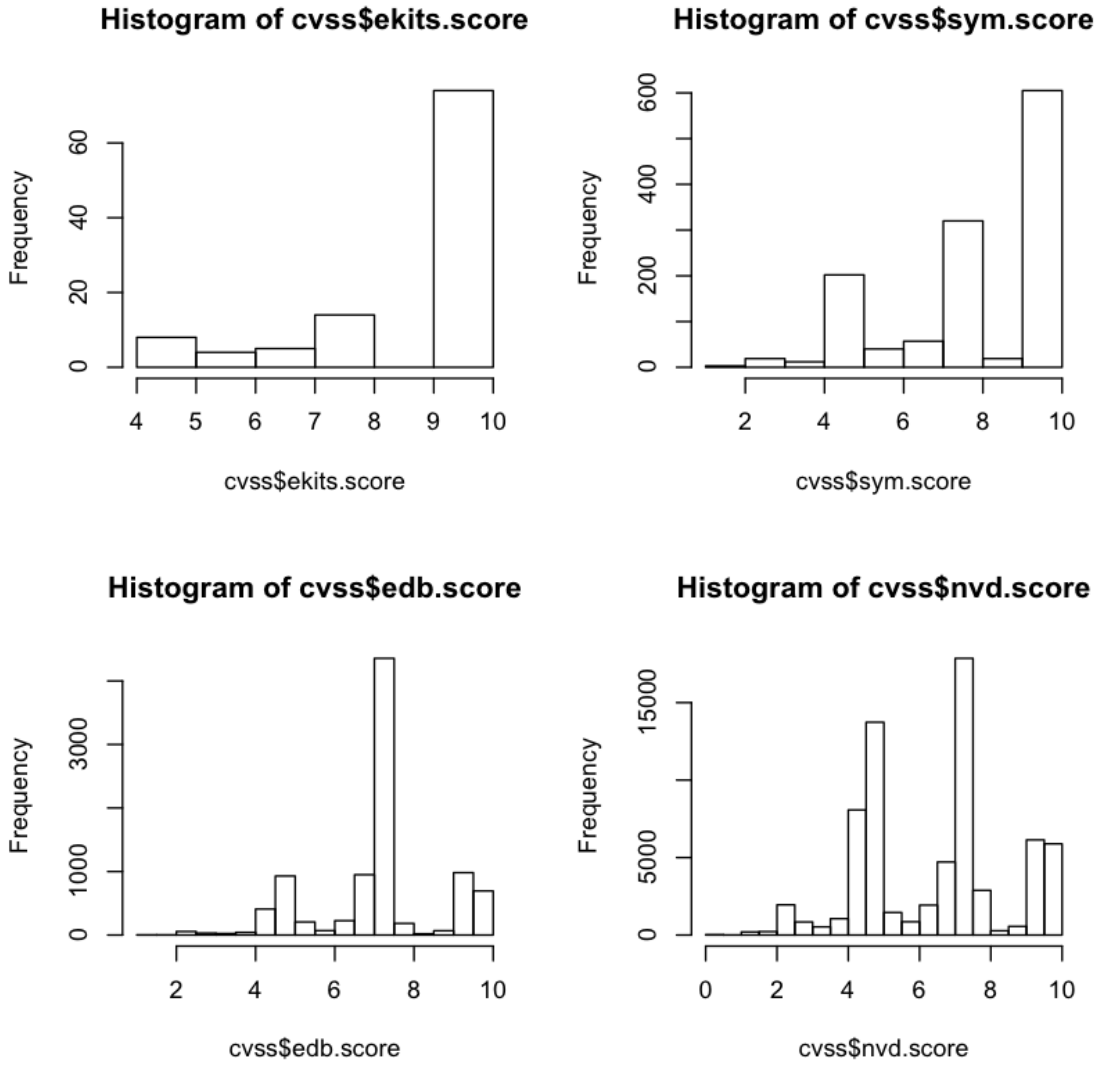
CVSS Study

- Remember: the SCAP protocol tells you: take a dataset of vulnerabilities, order vulnerabilities by CVSS.

- We therefore look at:

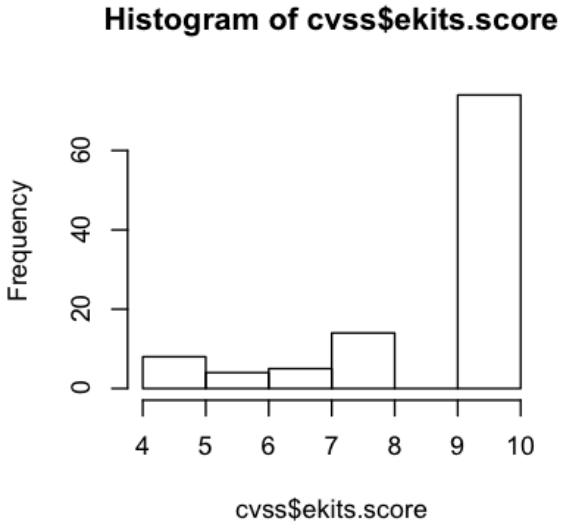
1. Distribution of CVSS scores per dataset
   - Are datasets different in terms of type of vulnerabilities?
2. VENN diagram of datasets and scores
   - Are datasets interesting in terms of attacks actually delivered by the bad guys?

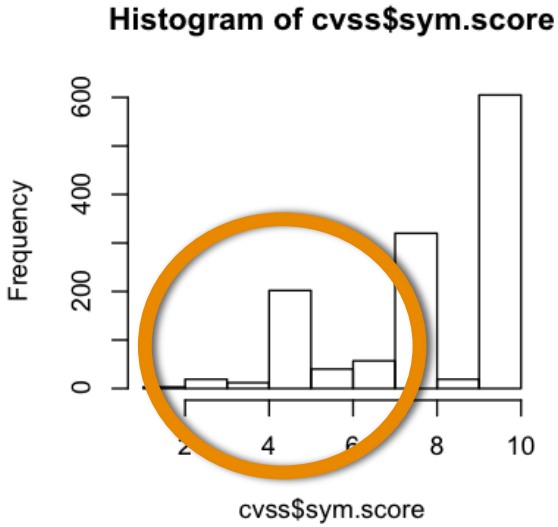
CVSS Distribution: HIST

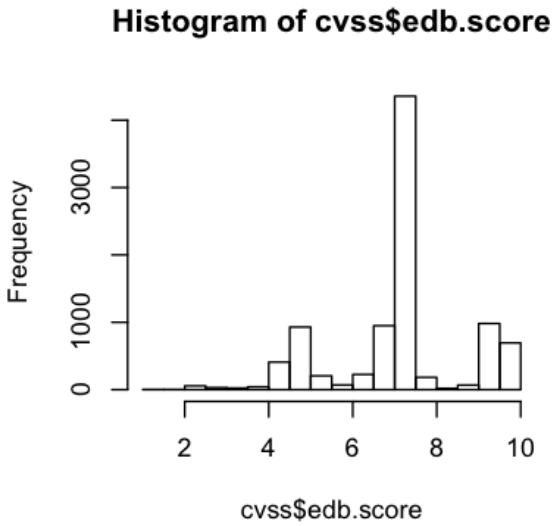


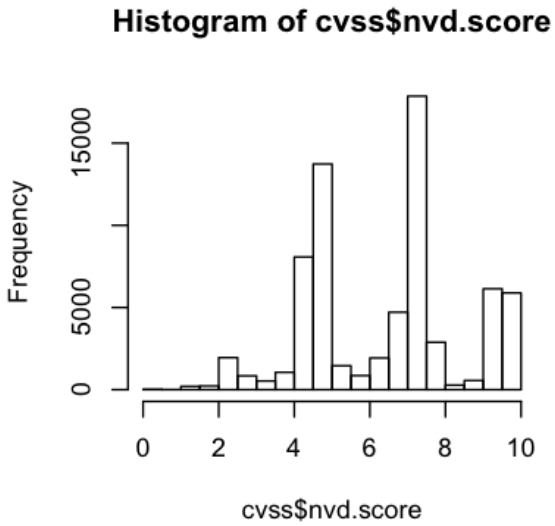
- LOW: CVSS <6
- MEDIUM: 6<CVSS<9
- HIGH: CVSS > 9

CVSS Distribution: HIST

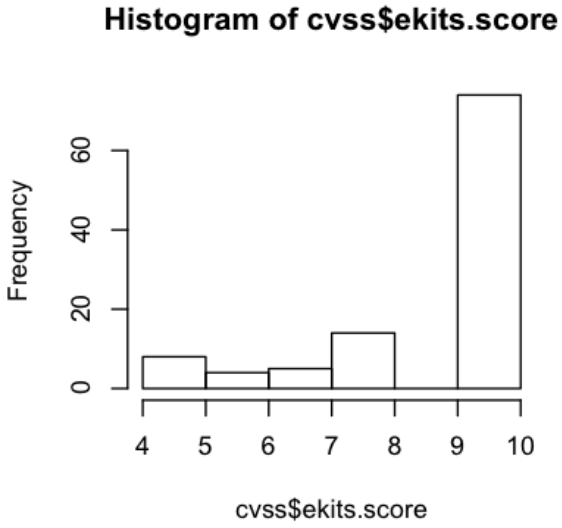


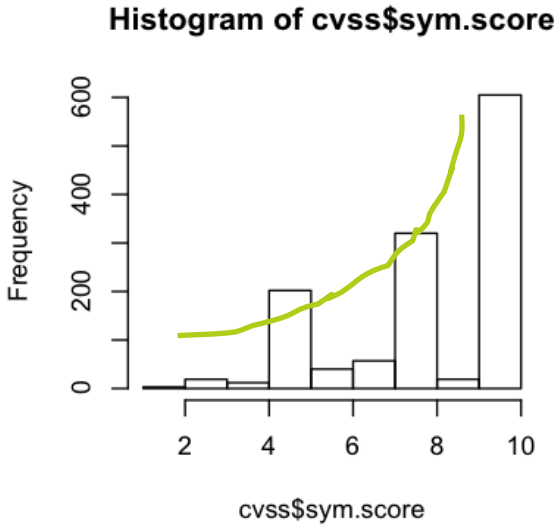


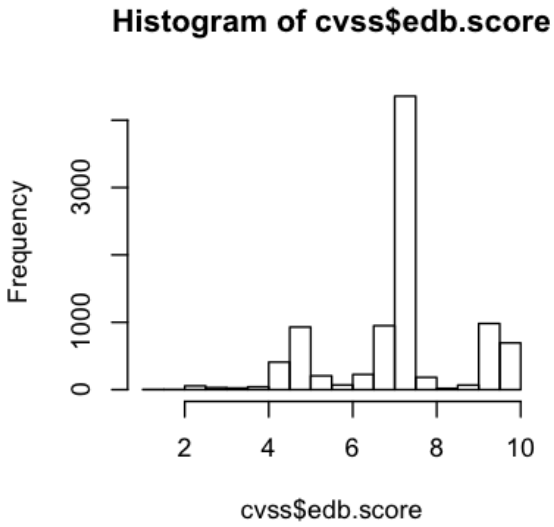


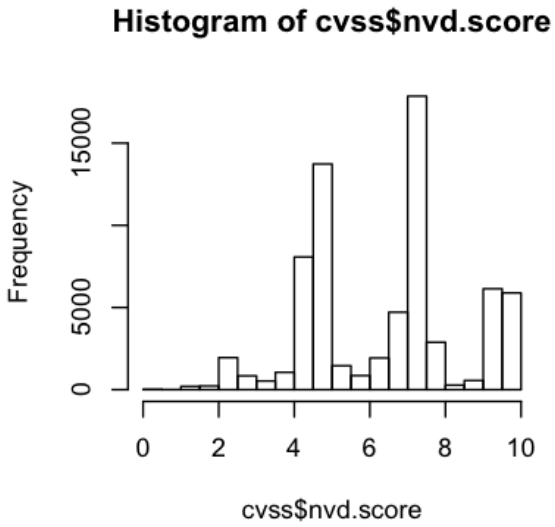


CVSS Distribution: HIST

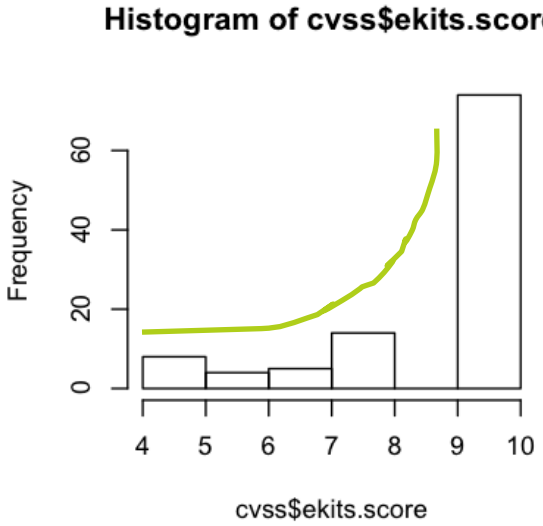


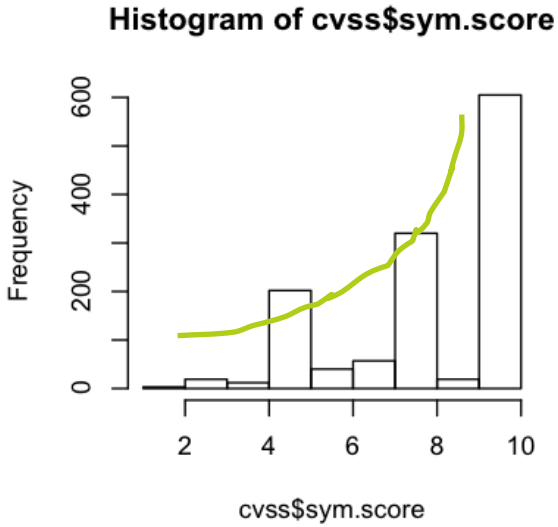


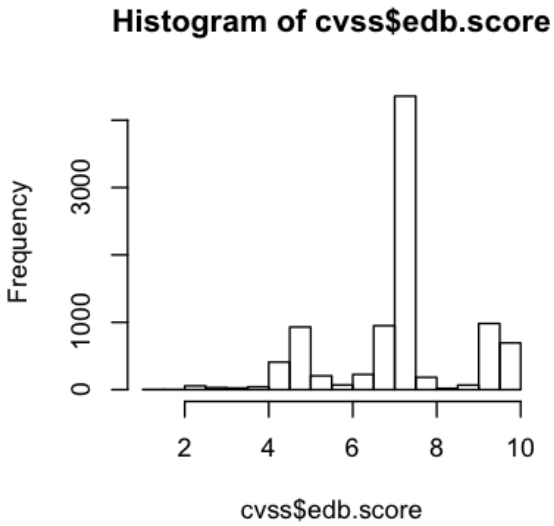


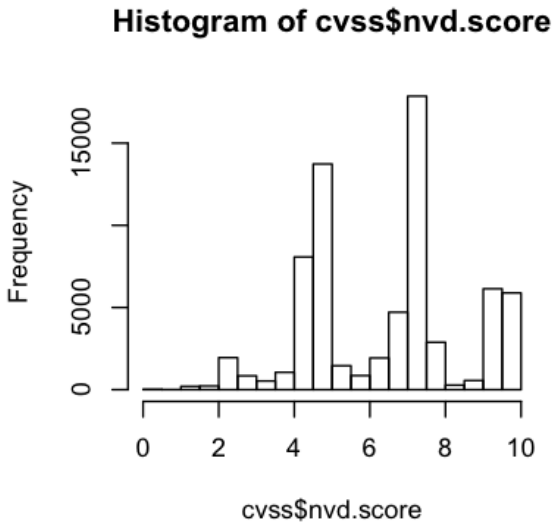


CVSS Distribution: HIST

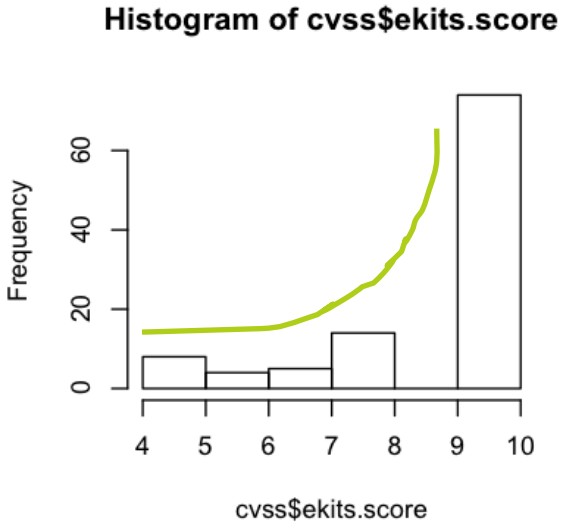


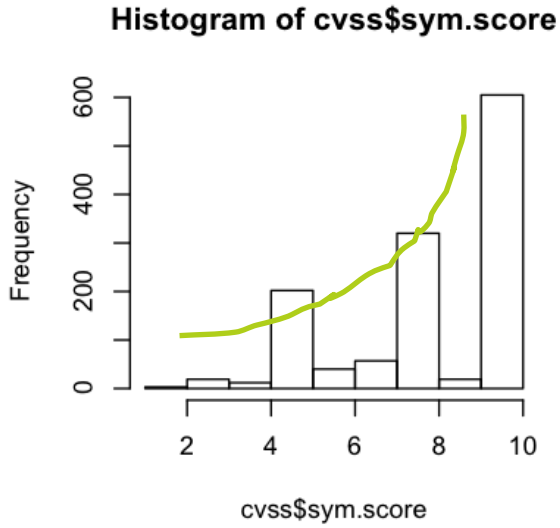


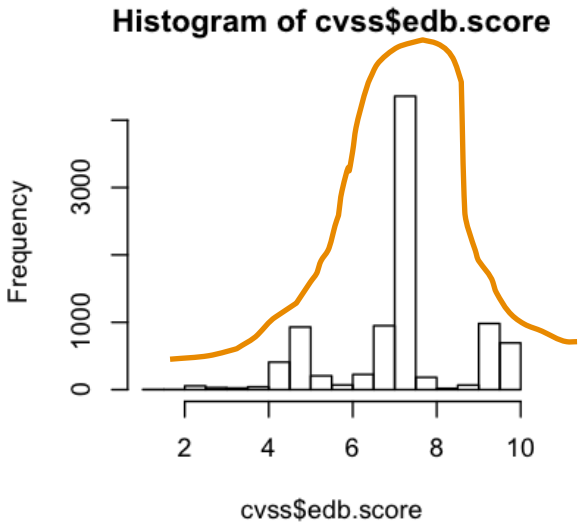


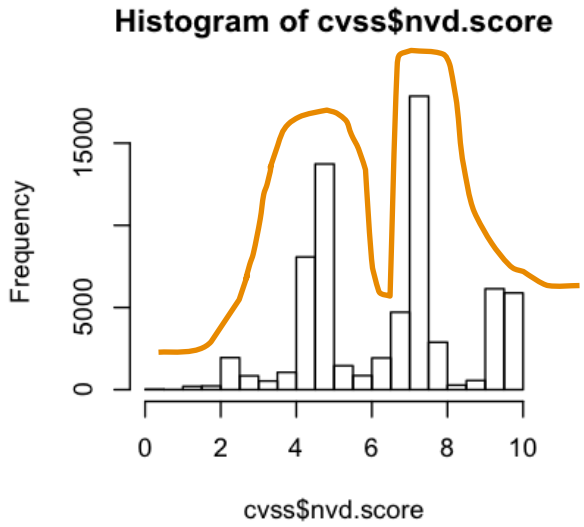


CVSS Distribution: HIST

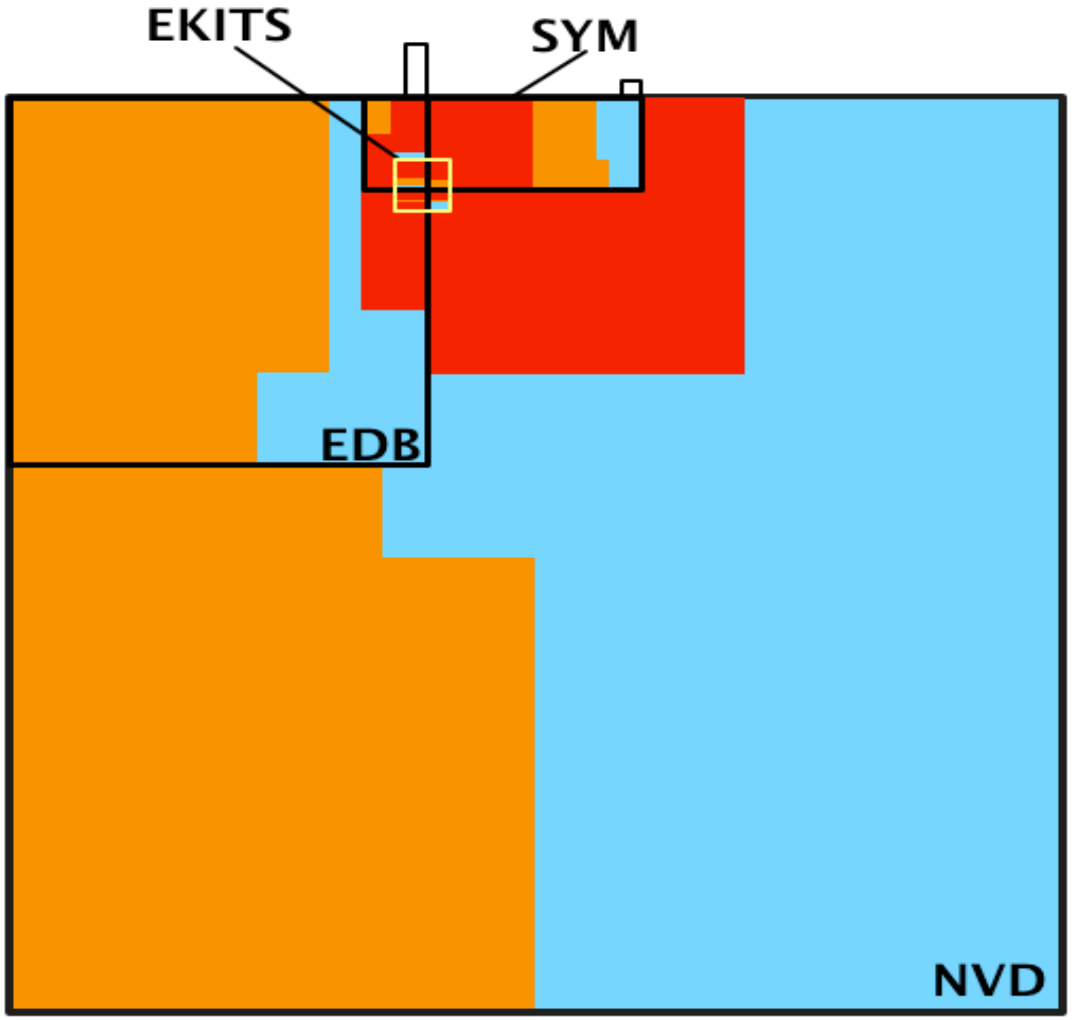








CVSS Distribution: VENN



LOW CVSS

MEDIUM CVSS

HIGH CVSS

Observational conclusions

- Attackers choose vulnerabilities autonomously:
  - They do not care about every vulnerability (NVD)
  - They do not care about every exploit (EDB)

- HIGH, MED+LOW score vulnerabilities are uniformly distributed in SYM dataset
- If you take NVD and fix all HIGH score vulnerabilities first [SCAP] you will:
  - Waste a lot of money patching all HIGH score vulnerabilities
  - Have addressed only 50% of final possible threats

What makes the CVSS so inaccurate?

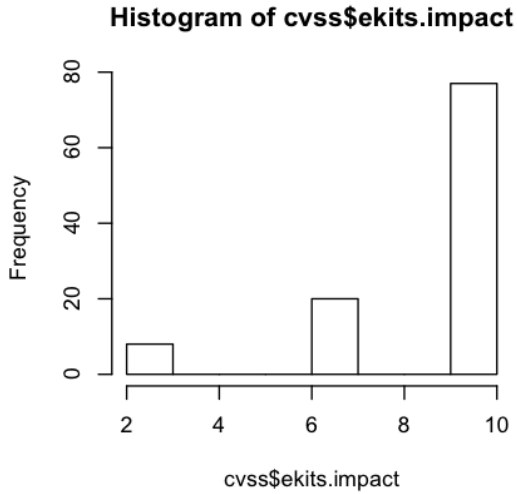
CVSS Metrics

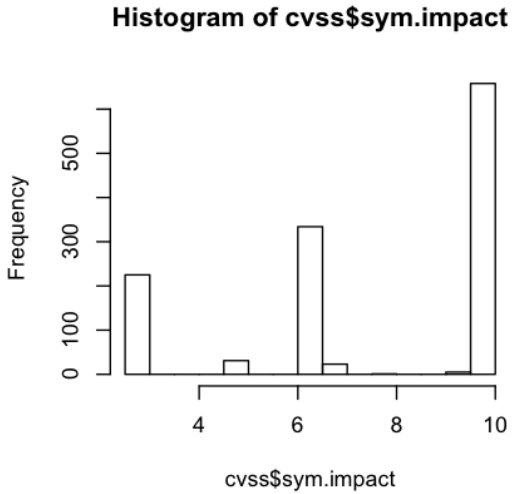
- CVSS measures risk in the form

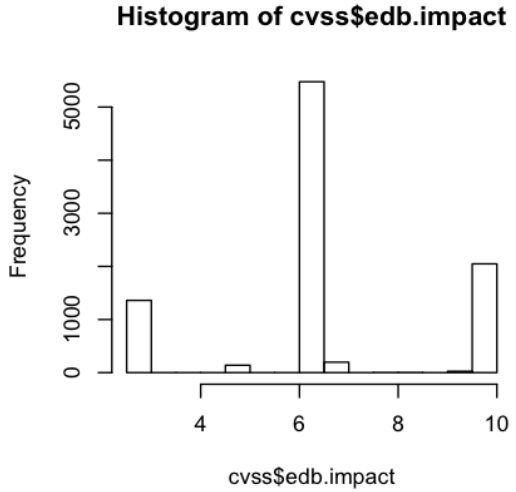
Risk = Impact x Likelihood

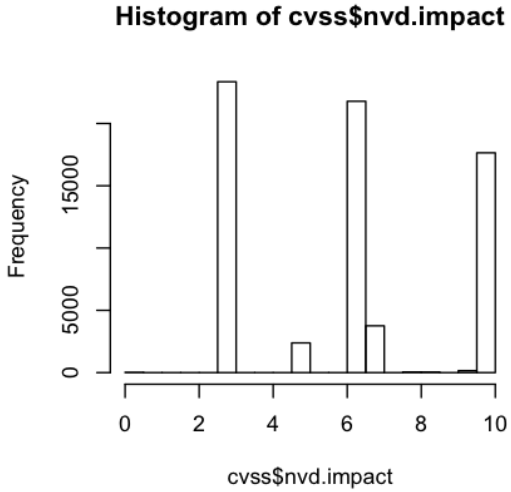
CVSS score = Impact x Exploitability

CVSS Metrics: Impact

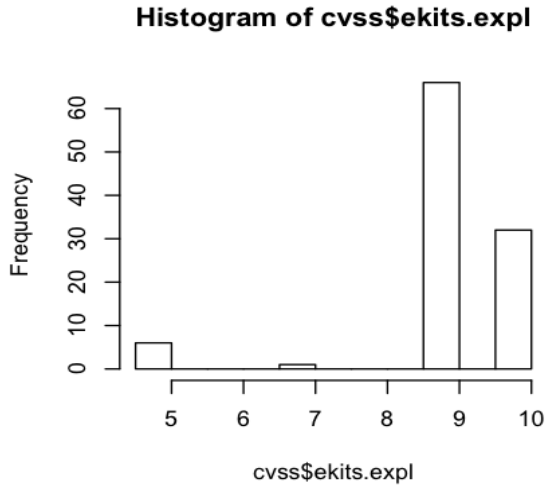


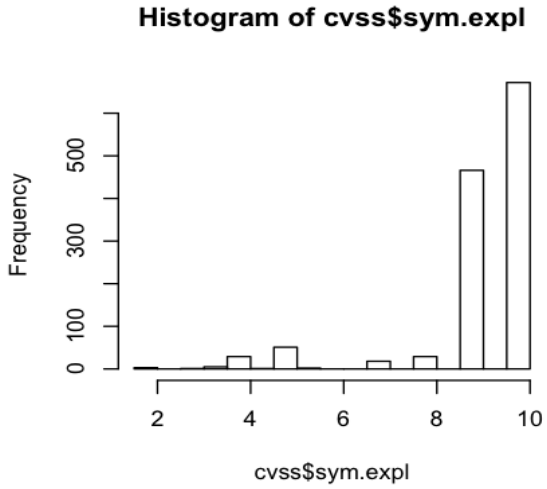


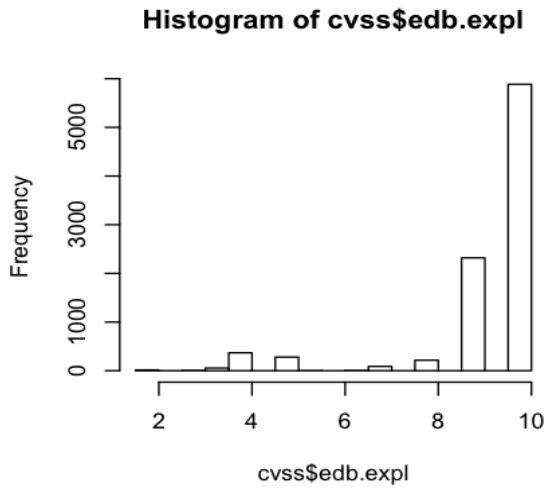


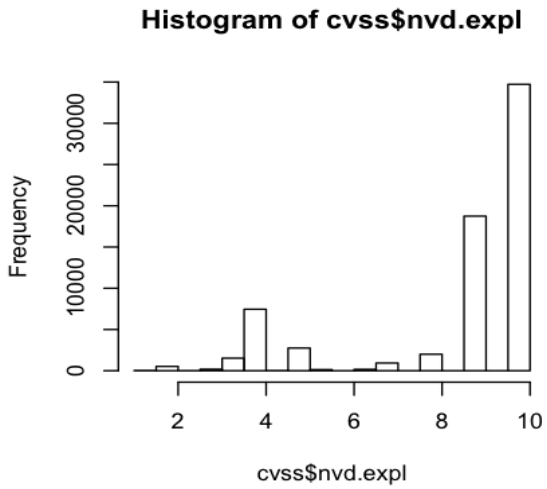


CVSS Metrics: Exploitability









CVSS Metrics: Exploitability explained

- Everything is exploitable → Exploitability is not an interesting variable at all!
  - Is actually a constant
- CVSS lacks of any real measure of likelihood
  - Based on "easiness to exploit"
    - Access Vector = All from Network VAR ≅ 0
    - Authentication = All None VAR ≅ 0
    - Access Complexity = Only interesting variable. VAR != 0
- Let's see what effects does this have to the final CVSS assessment

CVSS case controlled experiment

- Do smoking habits predict cancer? [Doll & Bradfor Hill, BMJ]
  - You can't ask people to start smoking so you can't run a controlled experiment
- Do high CVSS scores predict exploitation?
  - You can't attack users so you can't run a controlled experiment

CVSS case controlled experiment

Study	Cases	Controls (possible confounding variables)	Explanatory variable
Carcinoma of the lung	People with cancer	• Age • Sex • Location	• Smoke much Smoke some • Doesn't smoke
CVSS	Exploited vulnerabilities	• Access complexity • Access vector • Authentication • Impact type	• CVSS is HIGH • CVSS is LOW • Vuln is in {NVD,EDB,EKITS}

CVSS case controlled experiment

- CVSS Score+DB as a “medical test”

- Sensitivity → Pr(true positives)
  - You want to capture as many sick people as possible
- Specificity → Pr(true negatives)
  - You REALLY don’t want to cure people who don’t need it

CVSS Case Controlled Experiment

- Triple Blood Test Down Syndrome - Women aged 40+ [Kennard 1997]
  - Sensitivity: 69%
    - 31% of women carrying a fetus with Down syndrome will not be caught by the test
  - Specificity: 95%
    - only 5% of healthy pregnant women would be mislead by the test to undergo additional expensive or dangerous tests
  - Remember: most (but really a lot of) women have healthy pregnancies
- Prostate Serum Antigen - Men aged 50+ [Labrie 1992]
  - Sensitivity: 81%
  - Specificity: 90%

Security Rating as "Generate Panic" test

- Sensitivity: is High/Med CVSS good marker for v∈SYM?

Sensitivity = Pr(HIGH+MED | v in SYM)

- Specificity: is Low CVSS good marker for v∉SYM?

Specificity = Pr(LOW | v not in SYM)

Security Rating as "Generate Panic" test

DB	Sensitivity	Specificity
EKITS	89.17%	49.73%
EDB	98.14%	24.39%
NVD	89.70%	22.22%
3BT: Down Syndrome	69%	95%
PSA: Prostate Cancer	81%	90%

Security Rating as "Generate Panic" test - Explained

- **Sensitivity (+)**
  - CVSS is good in marking exploitation
- **Specificity (-)**
  - Peaks in NVD and EDB at less than 25%
  - 1 out of 4 non-exploited vulnerabilities are marked LOW
  - 3 out of 4 non-exploited vulnerabilities are marked HIGH
- Remember this is a controlled study:
  - We are looking only at vulnerabilities representative of SYM CVSS
- Let's assume linearity of cost for number of fixed vulnerabilities
- You are following US Governement SCAP Guidelines? -> You are spending up to 300% more money than you should

Ok, but is at least my risk decreasing?

- What really matters is change in relative probabilities

- Example = Usage of Safety Belts
  - Few people actually die in car crashes vs #crashes [Evans 1986]
  - Pr(Death x Safety Belt on) – Pr(Death x Safety Belt off)
  - 43% improvement of chances of survival
- Our Study = Patching High score vulnerabilities
  - Few vulnerabilities are actually exploited vs #vulns
  - Pr(Attack x CVSS High Patched) – Pr(Attack x CVSS Low Patched)
  - X% improvement of chances of NOT being attacked

Not really, no.

	Pr(H+M)-Pr(L)
EKIT	
vuln in SYM	+46.3%
vuln !in SYM	-47.28%
EDB	
vuln in SYM	+14.5%
vuln !in SYM	-14.49%
NVD	
vuln in SYM	+3.5%
vuln !in SYM	-3.46%

What does this mean?

- What the CIO really wants to know:
  - I read on the news that a "security researcher" exploited a vulnerability on X to do some bad stuff. Should I worry?
- You monitor the black markets and fix all HIGH CVSS vulnerabilities you find there?
  - Your risk of suffering from an attack from the black markets decreases by 46%
- You use EDB or NVD to know what exploits are out there, and fix all HIGH CVSS vulnerabilities?
  - Diminished risk: EDB = 14%; NVD = 3%.
  - Arguably a bad investment

Preliminary conclusions

- Where should we look for “real” exploits?
  - EDB, NVD are the wrong datasets
- Should the CIO do what SCAP protocol says?
  - No datasets shows high Specificity:
    - CVSS doesn’t rule out “un-interesting” vulns
    - Huge over-investment
- It may be possible to narrow down vulnerabilities the CIO should actually fix
  - Rule out 80% of risk = worth the update pain, measurable gain
  - We need better attacker model -> Research challange ahead

Questions

Thanks

What security researchers deliver

- Analysis of complete protection against a powerful adversary
- Attackers will target me in particular, intercept all my possible messages, exploit all my possible vulnerabilities, use all partners
  - Dolev, Schneier…
    - ➢ Fix all vulnerabilities or die

Vulnerabilities: most bad guys

- We are monitoring 90+ exploit kits on the markets
  - Automated infrastructure that monitors new kits, new CVEs, new posts by vendors
  - Last entry we detected: ~20 days ago: WhiteHole, 3 exploits, 1200$/month
    - Security press started talking about it 7+ days later
- New players pop up monthly if not weekly
  - 2-12 exploits each
  - Prices from 1000/year -> 2000/month
- Exploit as-a-service
  - Rent-an-infection-service
    - Pay in “traffic” or pay in dollars
  - “Clean” from AV
    - Symantec detects your exploit kit? Pay us, we’ll repack the attack
  - Free trials
    - I’m new, you don’t know me but I am good: try me

Vulnerabilities: reality according to data

- Google: automated attacks are 70% of final risk
- Symantec: 1.3k exploits out of 50k vulnerabilities
- Two scenarios:
  1. The bad guy wants you. Zero day exploit, not much you can do about it
  2. The bad guy just wants some. Will fish from the shoal, if you happen to be there and vulnerable you should have patched.

CVSS distribution explained

- They have different distributions!
  - EKITs sell mostly vulns with high scores
  - SYM see vulns with high scores and some with medium scores
    - Recall vuln in SYM → vuln used by bad guys
  - NVD and EDB have lots but really lots of vulns of totally uninteresting vulns
  - The population of exploited vulnerabilities (SYM) is different from NVD, EDB
- If you are using the NVD or EDB to assess your company status (eg SCAP) → Waste Money!
- CVSS scores tell something but not good enough
  - Only good for witch hunt - “Kill them all, God will recognize its brethren”

CVSS case controlled experiment

- Do smoking habits predict cancer?
  - Doll & Bradfor Hill, BMJ
  - You can't ask people to start smoking so you can't run a controlled experiment
- Case controlled study (Carcinoma)
  - Cases
    - people with lung cancer
  - Controls (Possible confounding variables)
    - Age, Sex, Social Status, Location
  - Explanatory variable
    - Smoking habit
  - For each of the cases select another person with the same values of the control variables

CVSS case controlled experiment

- Case controlled study (CVSS)
  - Cases
    - vulns with exploits in the wild (SYM/KASP)
  - Controls (Possible confounding variables)
    - Access vector, access complexity, authentication
  - Explanatory variables
    - CVSS Score, Database
- CVSS Score+DB as a “medical test”
  - Sensitivity → true positives vs all sick people
    - You want to capture as many sick people as possible
  - Specificity → true negatives vs all healthy people
    - You don’t want to cure people who don’t need it

Effects of removing High-risk Vulnerabilities

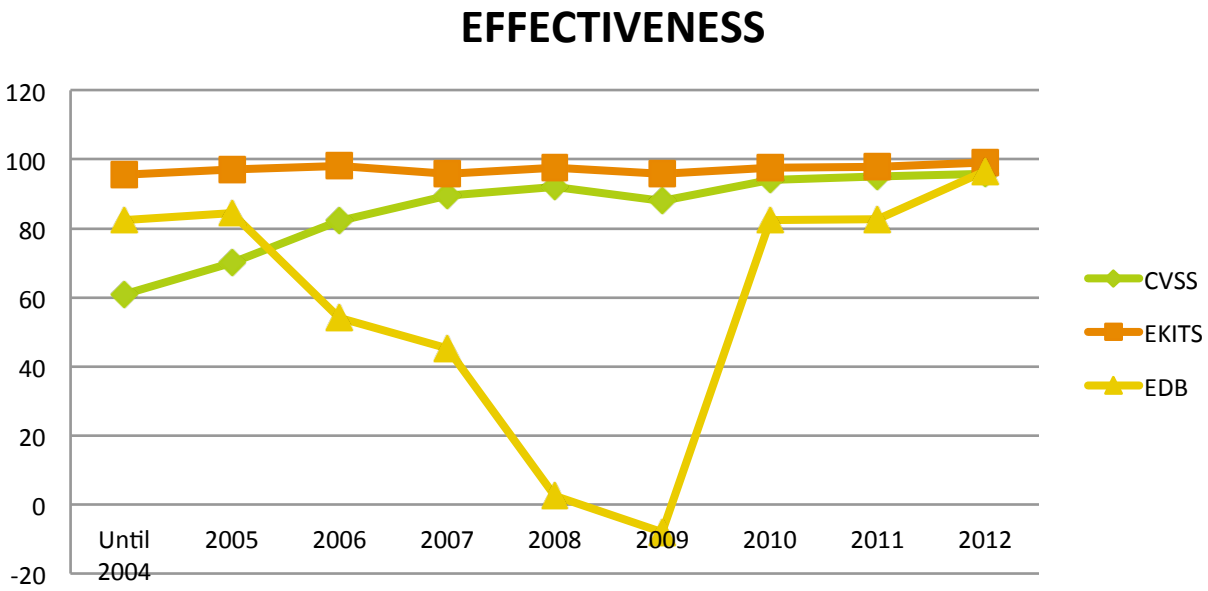
- We categorized our datasets per software type
  - 7 categories
  - BUSS(iness), PLUGIN, SERVER, WINDOWS, BROWSER…
- Test for Effectiveness considering as confounding variables:
  - Year
  - CATEGORY

Effects of Removing High-risk Vulnerabilities

- Classify vulnerabilities into two categories: High-risk vs. Low & Medium-risk
  - CVSS: CVSS (>9) vs. Low&Medium CVSS (<=9)
  - EDB & EKITS: In the dataset vs. Not in the dataset
- Based on Evans(1986), we calculate the effectiveness of removing high-risk vulnerabilities (i.e., w/ High CVSS, in EDB, and in EKITS) in reducing the risk of vulnerability exploitation in the wild.
- Effectiveness

Dataset	EFFECTIVENESS
CVSS	82.12 ± 2.06%
EDB	55.63 ± 5.22%
EKITS	96.80 ± 0.58%

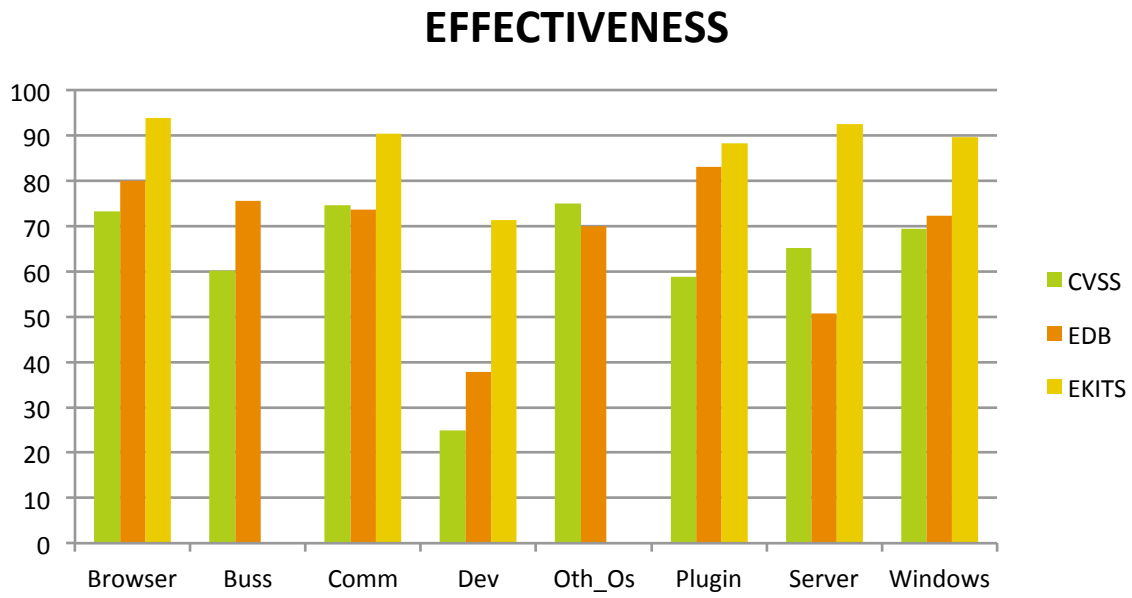
Effects of Removing High-risk Vulnerabilities (YEAR)



- Estimated Effectiveness (exponential model)

Dataset	EFFECTIVENESS
CVSS	56.77 ± 3.37%
EDB	36.81 ± 4.69%
EKITS	78.72 ± 3.39%

Effects of Removing High-risk Vulnerabilities (CATEGORY)



- Estimated Effectiveness

Dataset	EFFECTIVENESS
CVSS	39.80 ± 4.89%
EDB	42.11 ± 4.89%
EKITS	64.76 ± 6.03%

The Picture so Far

- What the CIO really wants to know:
  - I read on the news that a "security researcher" exploited a vulnerability on X to do some bad stuff.
  - Should we worry?
- The Answers…
  - A security researcher published a proof of concept exploit?
    - decline by 3-14% → delete email, life is too short
  - An exploit kit has marketed it and it has a CVSS high score?
    - decline by 46% → ask antivirus company or upgrade software, post a huge notice on the web site customers should update sw