Security Vulnerabilities: Risks from Discovery to Exploitation

Yashwant K. Malaiya
Colorado State University

CVSS: Common Vulnerability Scoring System

- How important is a specific vulnerability?
  - Essentially a risk measure
  - Vulnerabilities with highest scores need addressing quickly. Those with lowest scores are low priority.

- CVSS v1: *National Infrastructure Advisory Council* (NIAC), 2005
- CVSS v2: *Forum of Incident Response and Security Teams* (FIRST)
  - 2007
  - Still common
- CVSS V3: 2015
  - Getting common

CVSS: Common Vulnerability Scoring System

- Score formulas are based on metrics.
  - Metrics use table look-ups.
- Base Score uses metrics intrinsic to a vulnerability. Each official vulnerability (with a cve number) has a base score.
  - BaseScore = f(impact, exploitability)
  - Formulas designed to yield a value between 0 (lowest)-10 (highest). There is no formal derivation or justification for the formula.
  - Score used for prioritizing effort.

CVSS Scores

- BaseScore uses metrics intrinsic to a vulnerability. Each official vulnerability (with a cve number) has a base score.
  - Mandatory.
  - Impact based on values of Confidentiality, Integrity, and Availability (CIA) impact values.
- TemporalScore = f(BaseScore, Exploit, Remediation). Varies with time.
- EnvironmentalScore = f(metrics modified by required CIA levels for an application). Depends on the user environment.

CVSS v2.0 Base Score

Formula

- BaseScore = round_to_1_decimal(((0.6*Impact)+(0.4*Exploitability)-1.5)*f(Impact))
  - f(impact)= 0 if Impact=0, 1.176 otherwise
  - BaseScore ranges between 10-0
  - How did they come up with this?
  - No derivation, no validation, based on consensus in the committee based on member's expert opinions
- *Exploitability sub-score* - measure of Likelihood of exploitation of the vulnerability.
  - Range 0-10
- *Impact sub-score* - a measure of Impact.
  - Range 0-10

CVSS Base metric: Observation

- *CVSS Base Score* is a form of a risk measure.
- They could have computed *CVSS Base Score* by simply multiplying the *Exploitability* and the *Impact sub-scores*.
- It would result in a similar distribution of score with somewhat better resolution.
- *CVSS Base Score* for prioritizing vulnerabilities.
- V2 Base score
  - 7.0-10.0 **High** (V3: 7.0-8.9 High, 9.0-10 Critical)
  - 4.0-6.9 **Medium**
  - 0-3.9 **Low** (V3: 0.0 None, 0.1-3.9 Low)

CVSS 2.0 Exploitability Subscore

Exploitability = 20* AccessVector*AccessComplexity*Authentication

- AccessVector:
  - requires local access: 0.395
  - adjacent network accessible: 0.646
  - network accessible: 1.0
- AccessComplexity:
  - high: 0.35
  - medium: 0.61
  - low: 0.71
- Authentication
  - multiple instances of authentication: 0.45
  - requires single instance of authentication: 0.56
  - requires no authentication: 0.704

CVSS 2.0 Impact (C.I.A.) Subscore

Impact = 10.41*(1-(1-ConfImpact)*(1-IntegImpact)*(1-AvailImpact))

- ConfImpact
  - none: 0.0
  - partial: 0.275
  - complete: 0.660
- IntegImpact
  - none: 0.0
  - partial: 0.275
  - complete: 0.660
- AvailImpact
  - none: 0.0
  - partial: 0.275
  - complete: 0.660
- Weighted by required levels for specific environments for Environmental Score.

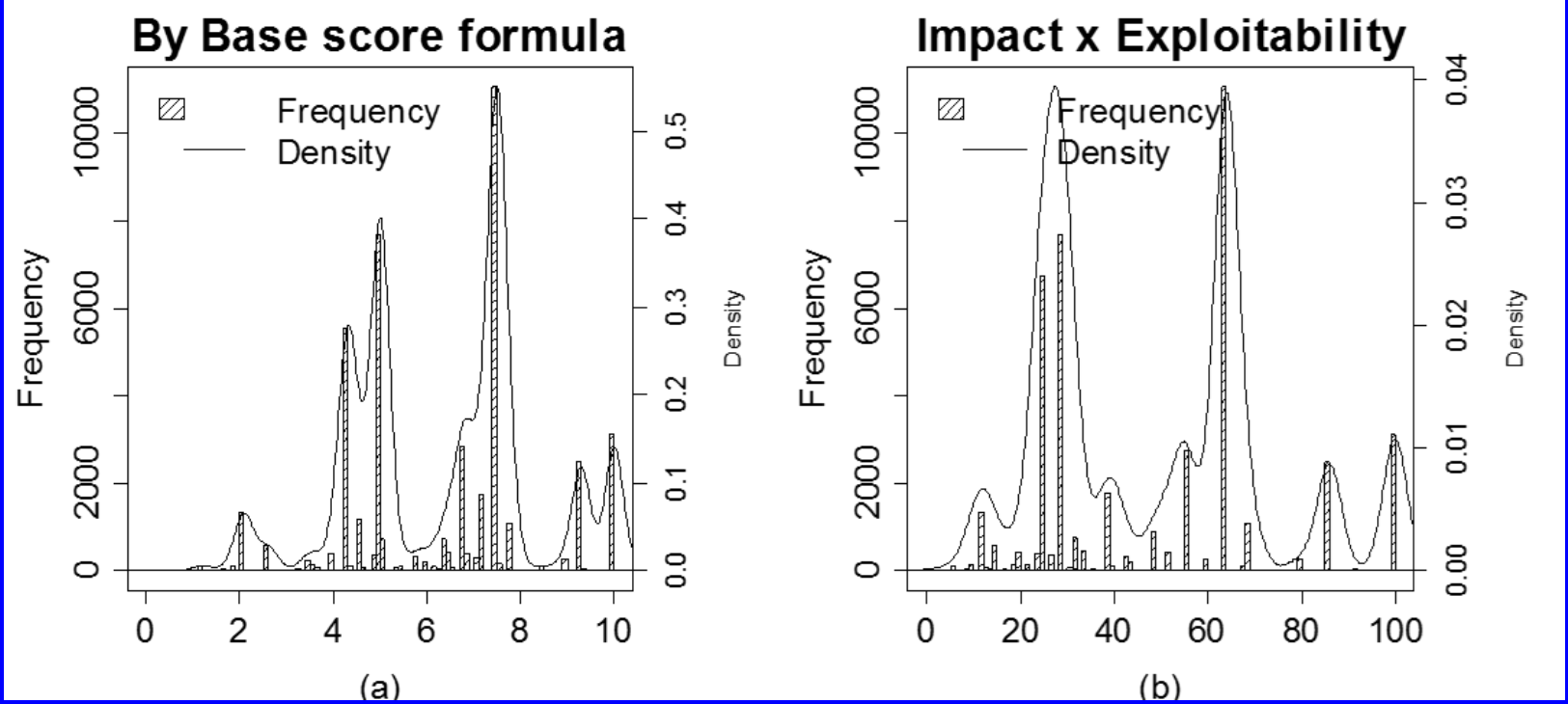
CVSS 2.0 other scores (not really used)

TemporalScore = round_to_1_decimal(BaseScore*Exploitability *RemediationLevel*ReportConfidence)

- Exploitability: Proven (1.0) to unproven (0.85)
- RemediationLevel: Official fix (0.85) to no fix (1.0)
- ReportConfidence: Confirmed (1.0) to unconfirmed (0.95)

EnvironmentalScore = round_to_1_decimal((AdjustedTemporal+ (10-AdjustedTemporal)*CollateralDamagePotential)*TargetDistribution)

Distribution of Base score



	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Combinations
(a)	0	5	6.8	6.341	7.5	10	63
(b)	0	29	49	48.59	64	100	112

NVD on Jan 2011 (44615 vuln.)

H. Joh and Y. K. Malaiya, "Defining and Assessing Quantitative Security Risk Measures Using Vulnerability Lifecycle and CVSS Metrics," SAM'11, The 2011 International Conference on Security and Management, pp.10-16, 2011

Has CVSS worked?

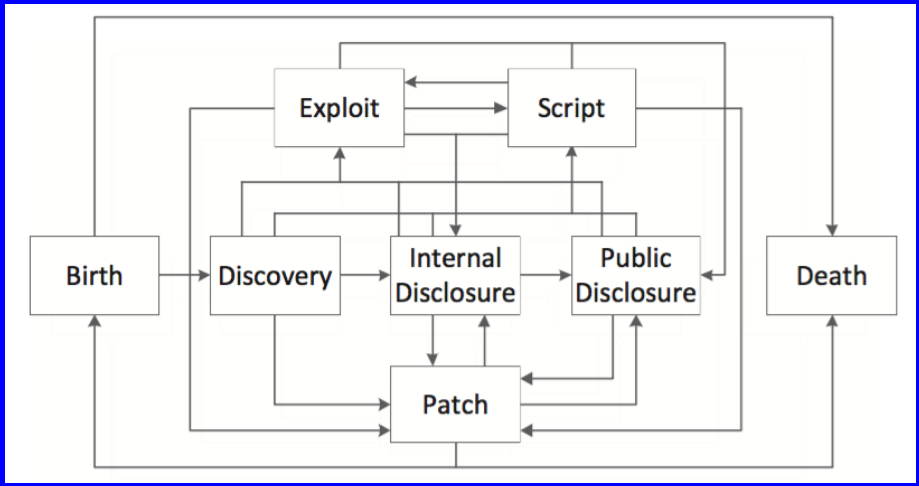
- Windows 7 Correlation among
  - CVSS Exploitability
  - Microsoft Exploitability metric
  - Presence of actual exploits
- No significant correlation found.

Variables	Exploit Existence	MS-EXP	CVSS-EXP
Exploit Existence	1	-0.078	-0.146
MS-EXP	**-0.078**	1	-0.116
CVSS-EXP	**-0.146**	-0.116	1

A. Younis and Y.K. Malaiya, "Comparing and Evaluating CVSS Base Metrics and Microsoft Rating System", The 2015 IEEE Int. Conf. on Software Quality, Reliability and Security, pp. 252-261

Likelihood of Individual Vulnerabilities Discovery

- **Ease of discovery**
  - Human factor (skills, time, effort, etc.), Discovery technique, Time
- Time:



- Apache HTTP server
- CVE-2012-0031,
- Published (01/18/2012)
- First appeared V. 1.3.0→1998-06-06
- 13.5 years

Time to Discovery = Discovery Time Date – First Effected version Release Date

Correlation: Access complexity vs Time to Discover

AC vs Time

- **AC= Low**

Min.	1st Qu.	**Median**	Mean	3rd Qu.	Max.
0.100	0.900	**2.000**	3.338	4.500	18.000

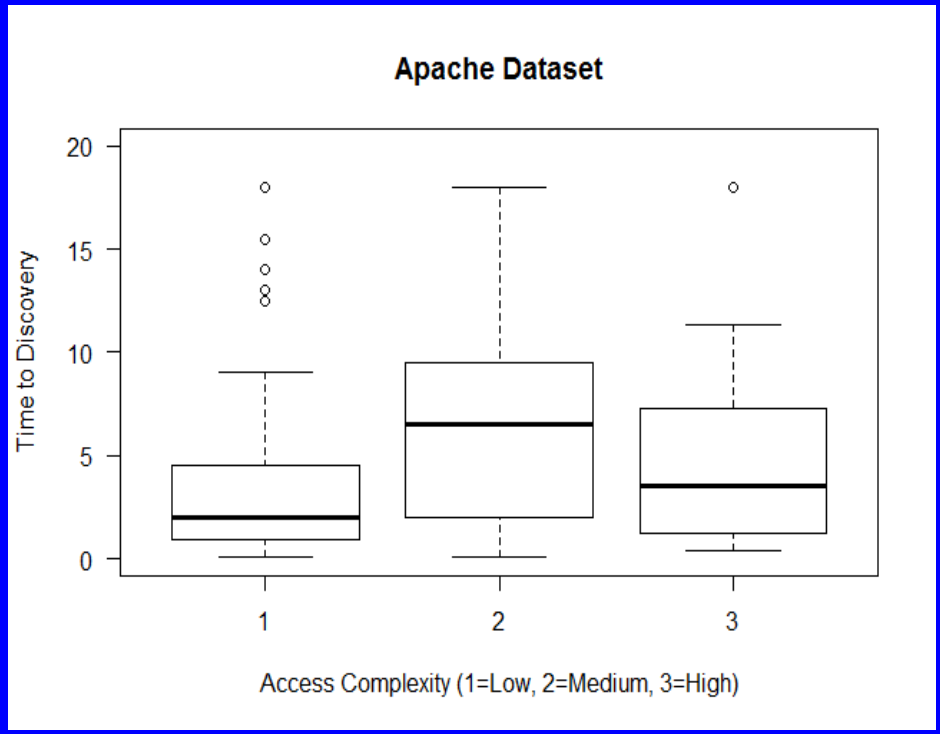
- **AC= Medium**

Min.	1st Qu.	**Median**	Mean	3rd Qu.	Max.
0.100	2.000	**6.500**	6.819	9.500	18.000

- **AC= High (very few points)**

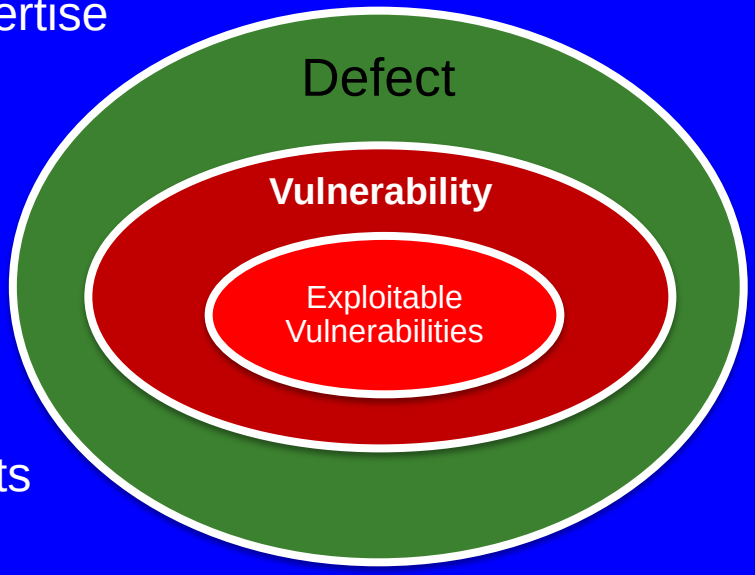
Min.	1st Qu.	**Median**	Mean	3rd Qu.	Max.
0.400	1.350	**3.500**	5.208	7.125	18.000

- **Some correlation between Access Complexity and Time to Discover**



Characterizing Vulnerability with Exploits

- **1 to 5 %** of defects are vulnerabilities.
- Finding vulnerabilities can take considerable expertise and effort.
- Out of 49599 vulnerabilities reported by NVD, **2.10% have an exploit**.
- A vulnerability with an exploit written for it presents more risk.

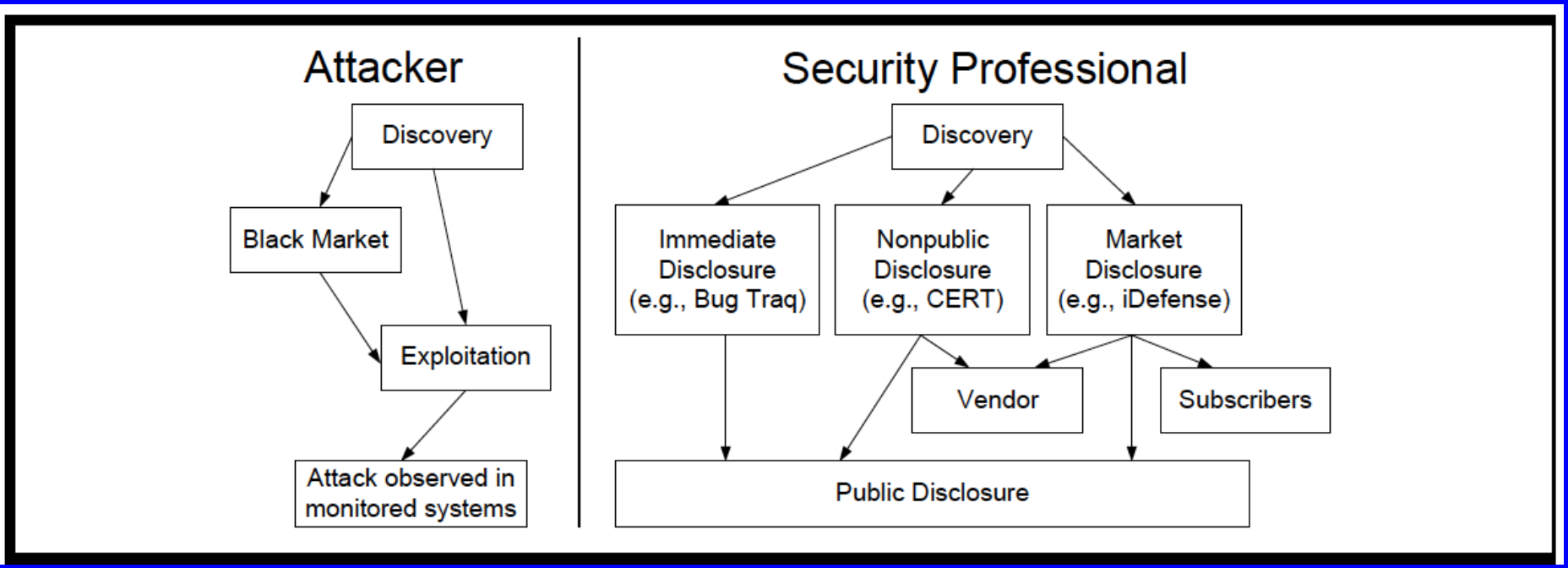


- **What characterizes a vulnerability having an exploit?** Small sloc

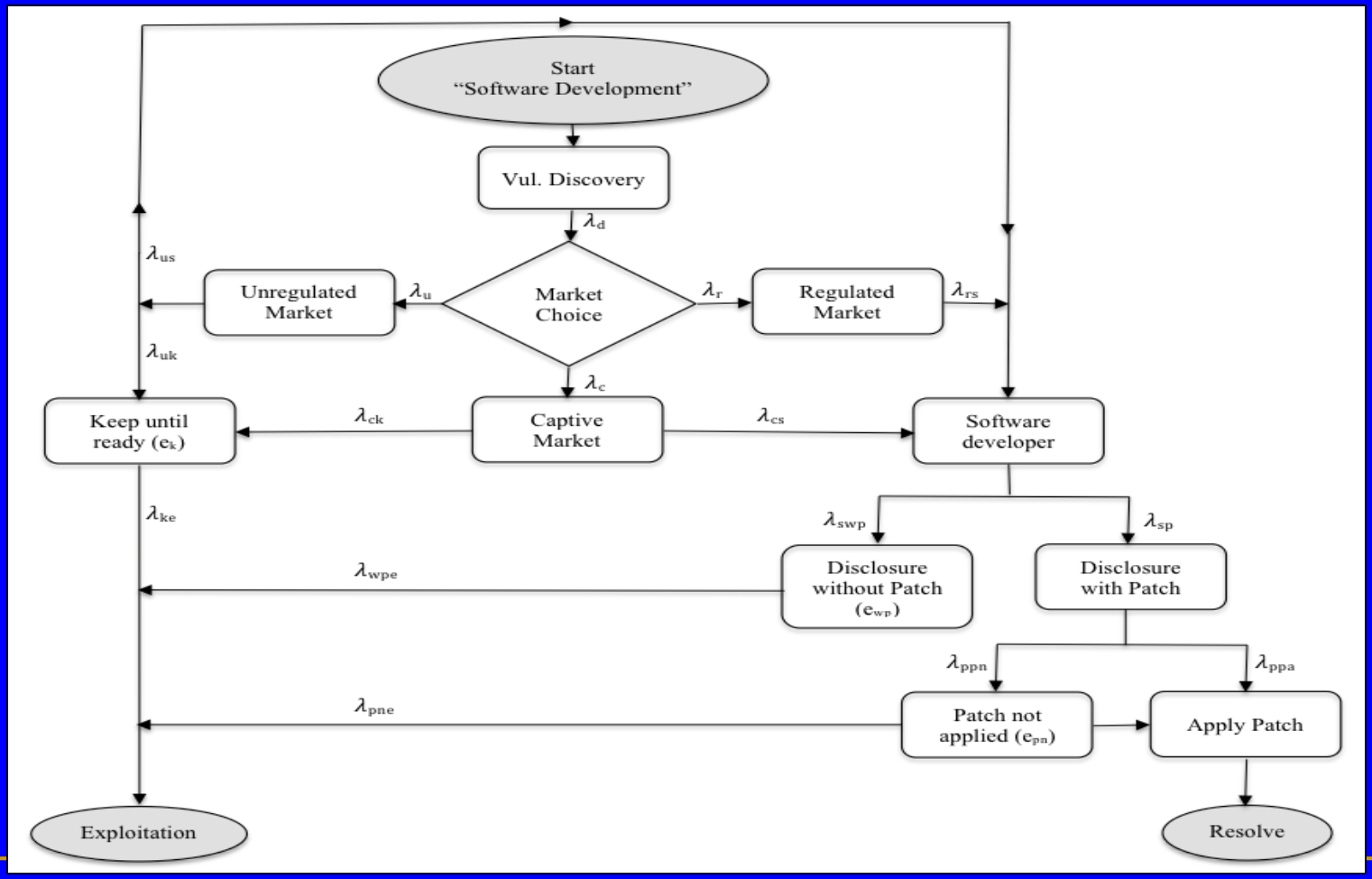
Vulnerability	In-Degree	Out-Degree	CountPath	ND	CYC	Fan-In	No of Invocation	SLOC	Exploit Existence
CVE-2009-1891	1	9	9000	6	68	45	2	211	NEE
CVE-2010-0010	4	9	145	4	11	16	4	38	EE
CVE-2013-1896	26	5	8	1	5	37	3	29	EE

Awad Youngish, Yashwant K. Malaiya, Charles Anderson, and Indrajit Ray. "**To Fear or Not to Fear That is the Question: Code Characteristics of a Vulnerable Function with an Existing Exploit**". Proceedings of the Sixth ACM on Conference on Data and Application Security and Privacy (CODASPY), 2016, pp. 97-104.

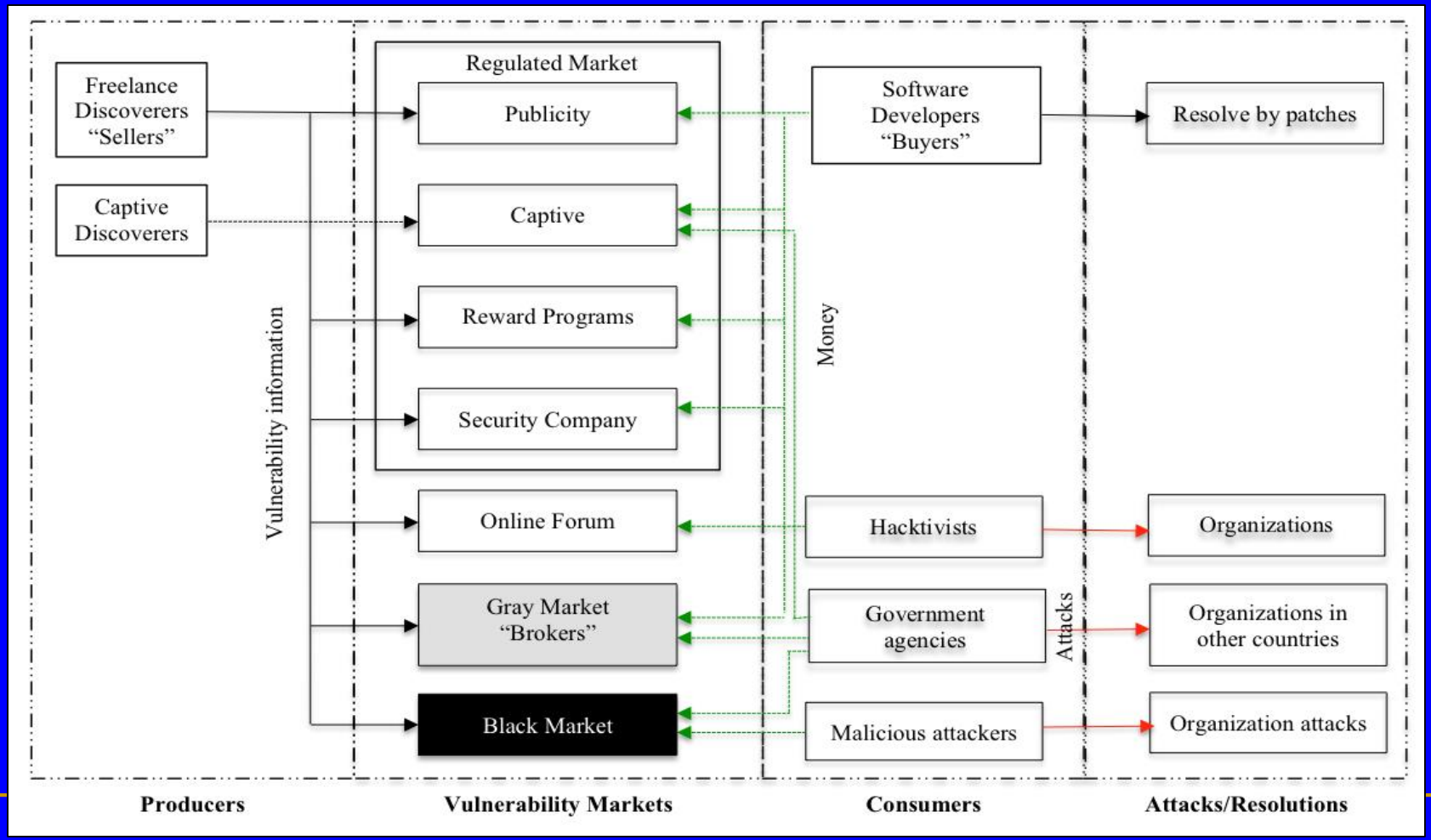
Vulnerabilities & Money



Vulnerability flow through markets



Types of Vulnerability Markets



Vulnerability Reward Programs (VPR)

- VRPs decreases the probability of an attacker acquiring a vulnerability and that reduces the likelihood of vulnerabilities discovery and exploitation.
- Vulnerabilities with a high CVSS scores and have no exploits or attacks may been explained by the impact of VRPs on vulnerabilities exploitation.
- We found significant correlation.

Spearman Correlation between CVSS Base score and VRPs Rating System			
	Correlation	**CVSS Scores before clustering**	**CVSS Scores after clustering**
Firefox	Value	0.65	0.47
	P-value	< 0.0001	< 0.0001
Chrome	Value	0.53	0.59
	P-value	< 0.0001	< 0.0001

Awad Youngish, Yashwant K. Malaiya, and Indrajit Ray. "**Evaluating CVSS Base Score Using Vulnerability Rewards Programs**". The proceedings of the 31st International Conference on Systems Security and Privacy Protection (IFIP SEC 2016).

Measuring Impact

- Technical impact in terms of Confidentiality, Integrity, Availability: CVSS Impact factor, not validated
- Methods for computing business impact are still evolving
- Widely different estimates of financial impacts of recent security breaches
- Our effort is continuing

CVSS V3.0/3.1: What is new

- **User Interaction:**
  - Successful exploit requires a user to take some action to be exploited.
- **Scope**
  - The ability for a vulnerability in one software component to impact resources beyond its means, or privileges
- Formulas and scales somewhat different
- On-line calculators available: v2.0 v3.0

CVSS Example: cvedetails.com

Try

- https://www.cvedetails.com/cve/CVE-2018-1000006/
- **Vulnerability Details : CVE-2018-1000006 (1 Metasploit modules)**
- CVSS Scores & Vulnerability Types
- Products Affected By CVE-2018-1000006
- Number Of Affected Versions By Product
- References For CVE-2018-1000006
- Vulnerability Conditions

Security Breach Cost Models

- **Ponemon Institute**
  - Founded in 2002 by Larry Ponemon and Susan Jayson
  - conducts independent research on data protection
  - Collaborates with several large organizations and publishes annual reports
- **NetDiligence**
  - Privately-held cyber risk assessment and data breach services company.
  - Since 2001, NetDiligence has conducted thousands of enterprise-level cyber risk assessments for a broad variety of organizations
  - NetDiligence services are used by leading cyber liability insurers in the U.S. and U.K.
- **Ponemon assisted models, sponsored by**
  - Symantac (2010),
  - Megapath (2013), and
  - IBM (2014)
- **NetDiligence Model**
  - Hub International calculator (2012) and
  - contributed to the Verizon report

Security Breach Cost Metrics

Total Cost of a Breach =

Incident investigation cost

+ Customer Notification/crisis management cost

+ Regulatory and industry sanctions cost

+ Class action lawsuit cost

$$\boldsymbol{Cost\ per\ Record} = \frac{Total\ cost\ of\ breach}{number\ of\ affected\ records}$$

Cost Models: Investigations

- **The Ponemon Institute and NetDiligence data/models**
  - They used proprietary data available to them.
  - They derived computational models based on their data ("calculators").
  - Large number of factors, considerable variation in factors considered.
- **Objective of study by Algarni and Malaiya**
  - Identify the major factors that are significant
  - Build models for the factors identified.
- **Approach**
  - regenerate data using the computational engines by providing a large number of input combinations.
  - Identified and removed the factors that emerged as non-significant.
  - Developed systematic computational models.

Cost Models: Investigations

- **The Ponemon Institute and NetDiligence data/models**
  - They used proprietary data available to them.
  - They derived computational models based on their data ("calculators").
  - Large number of factors, considerable variation in factors considered.
- **Objective of study by Algarni and Malaiya**
  - Identify the major factors that are significant
  - Build models for the factors identified.
- **Approach**
  - regenerate data using the computational engines by providing a large number of input combinations.
  - Identified and removed the factors that emerged as non-significant.
  - Developed systematic computational models.

A consolidated approach for estimation of data security breach costs, AM Algarni, YK Malaiya 2016 2nd International Conference on Information Management (ICIM), 26-39

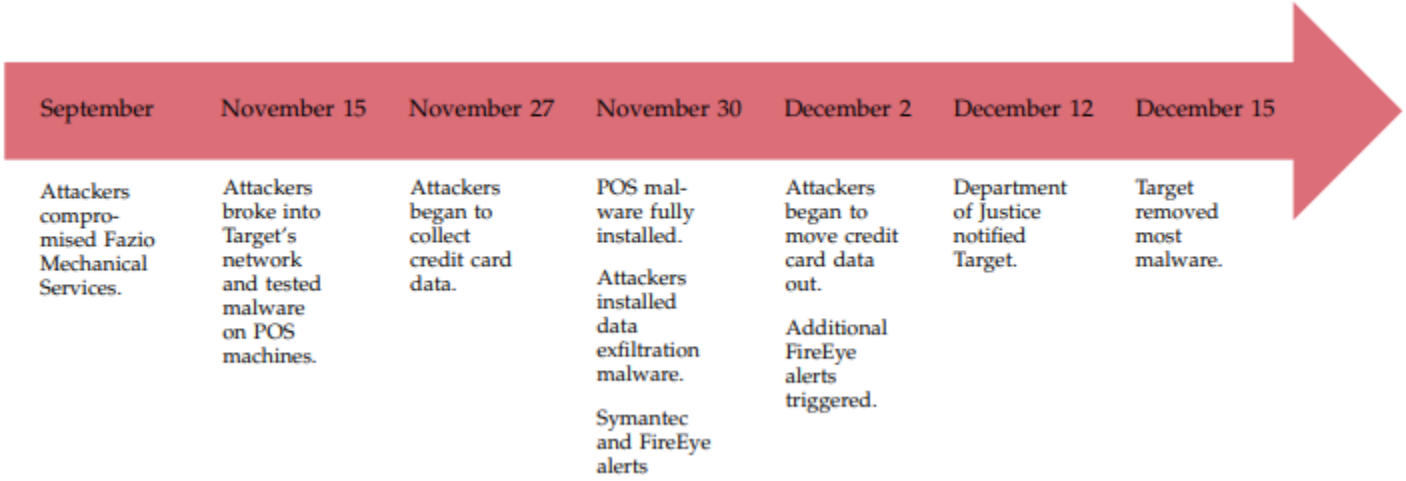
Significant Factors impacting Cost and Probability

Classification	Significant Factors	Source
Total number of affected records	Total number of affected records?	All
Type of data breaches	What is your organization's industry classification?	Symantec & IBM
	What types of information do your employees handle?	Symantec & IBM
Incident investigation cost	Data is in a centralized system/location?	Hub Int'l
	Actual fraud is expected already?	Hub Int'l
	Federal class action lawsuit filed?	Hub Int'l
	What do you think is the most likely cause of a data breach?	Symantec & IBM
	Is sensitive data encrypted on all laptops or removable storage?	Symantec & IBM
	How long does the business keep/retain sensitive information pertaining to employees, customers, and patients?	IDT911
	What best describes your organization's privacy and data protection program?	Symantec & IBM
Crisis management cost	Number of Years for credit monitoring?	Hub Int'l
	What is the global headcount of your organization?	Symantec & IBM
	Is your organization's business continuity management team involved in the data breach incident response process?	IBM
Regulatory and sanction cost	Is PCI compliance an issue?	Hub Int'l
Lawsuit cost	Actual fraud is expected already?	Hub Int'l
	Federal class action lawsuit filed?	Hub Int'l

A consolidated approach for estimation of data security breach costs, AM Algarni, YK Malaiya 2016 2nd International Conference on Information Management (ICIM), 26-39

Target data breach (2013)

- Target Corporation's network
- Breach Dates: Between November 27 and December 18, 2013
  - Announced Dec 19, 2013 to media (Dec 18 KrebsOnSecurity, WSJ)
  - second largest credit and debit card breach after the TJX breach in 2007.
  - 40 million credit and debit card numbers and 70 million records of personal information were stolen.
  - It cost credit card unions over two hundred million dollars for just reissuing cards.
  - Wildly different cost estimates by experts, up to a billion.



Xiaokui Shu, Ke Tian, Andrew Ciambrone, and Danfeng Yao. Breaking the Target: An Analysis of Target Data Breach and Lessons Learned. CoRR, abs/1701.04940, 2017

Target data breach (2013)

- TGT Price chart (Yahoo Finance)

Note:

Colorado State University

TARGET DATA BREACH ACTUAL REPORTED COSTS

Years	Gross Expenses	Insurance receivable	Net Expenses (before tax deductions)	Net Expenses (after tax deductions)
2013	$61m	$44m	$17m	$11m
2014	$191m	$46m	$145 m	$94m
2015	N/A	N/A	$39	$28
Total	$252m	$90m	$201m	$133m
Raw cost per card= $6.30 (40 million cards affected)				

A consolidated approach for estimation of data security breach costs, AM Algarni, YK Malaiya 2016 2nd International Conference on Information Management (ICIM), 26-39

Home Depot Data Breach Actual reported Costs

- September 8th, 2014, Home Depot released a statement indicating that its payment card systems were breached.
- The data breach occurred from a sophisticated custom-built malware program installed on Home Depot's payment system network using a third-party vendor's login credentials.

Case Study: The Home Depot Data Breach, Brett Hawkins, 2015

Home Depot Data Breach Actual reported Costs

Years	Gross Expenses	Insurance receivable	Net Expenses (before tax deductions)	Net Expenses (after tax deductions)
Q3, 2014	$43m	$15m	$28m	N/A
Q4, 2014	$20m	$15m	$5m	N/A
Total	$63m	$30m	$33m	N/A
Raw cost per card= $1.13 (56 million cards affected)				

A consolidated approach for estimation of data security breach costs, AM Algarni, YK Malaiya 2016 2nd International Conference on Information Management (ICIM), 26-39

Is there an average cost per record?

- Using averages make sense, at least for initial estimates
- The **law of large numbers**:
  - sample size grows, its mean gets closer to the average of the whole population.
- The Flaw of Averages:
  - $2 billion in property damage in North Dakota.
  - In 1997, the U.S. Weather Service forecast that North Dakota's rising Red River would crest at 49 feet.
  - Officials in Grand Forks made flood management plans based on this single figure.
  - The river crested above 50 feet, breaching the dikes, and unleashing a flood that forced 50,000 people from their homes.

The Flaw of Averages, Sam Savage, Harvard Business Review, Nov. 2002

Ponemon: 2015 Cost of Data Breach in US

Partial costs	Avg. cost per breach	Avg. cost per record
Detection & escalation (includes investigation and crisis management)	$610,000	$21.73
Notification (includes notification and determination of regulatory)	$560,000	$19.95
Ex-post response (includes regulatory and lawsuit)	$1,640,000	$58.43
Lost business (includes reputation loss)	$3,720,000	$132.53
Total costs	$6,530,000	$217
Average number of records= 28070		

Average Cost per record: Hub Int.

Partial costs	Avg. cost per record for CC	Avg. cost per record for PHI&SSN
Incident investigation	$1.15	$1.64
Crisis management	$3.52	$4.57
Sanctions	$0.81	$0.81
Lawsuit	$7.09	$1.56
Total costs	$12.57	$8.58

Credit cards, Personal Health Information, SSN

2015 Cost of Data Breach in United States, Ponemon Institute

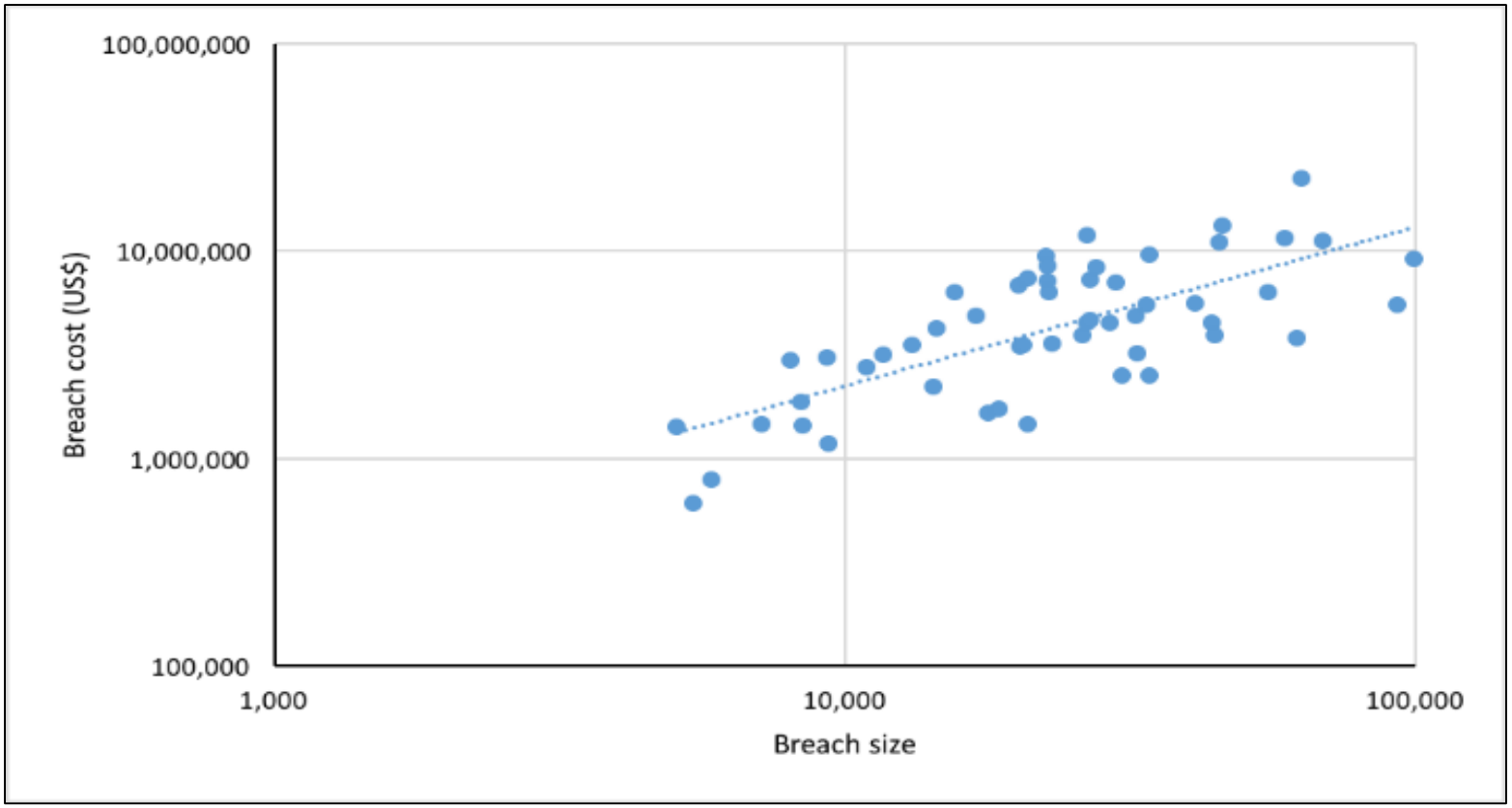
Average Cost per record

- What is the right number for *average cost per record*?
  - $217 Ponemon?
  - $8-$13 Hub International?
  - $0.58 Verizon?
- Controversy

Ken Spinner, Data breach cost estimates get it wrong: What you need to know.

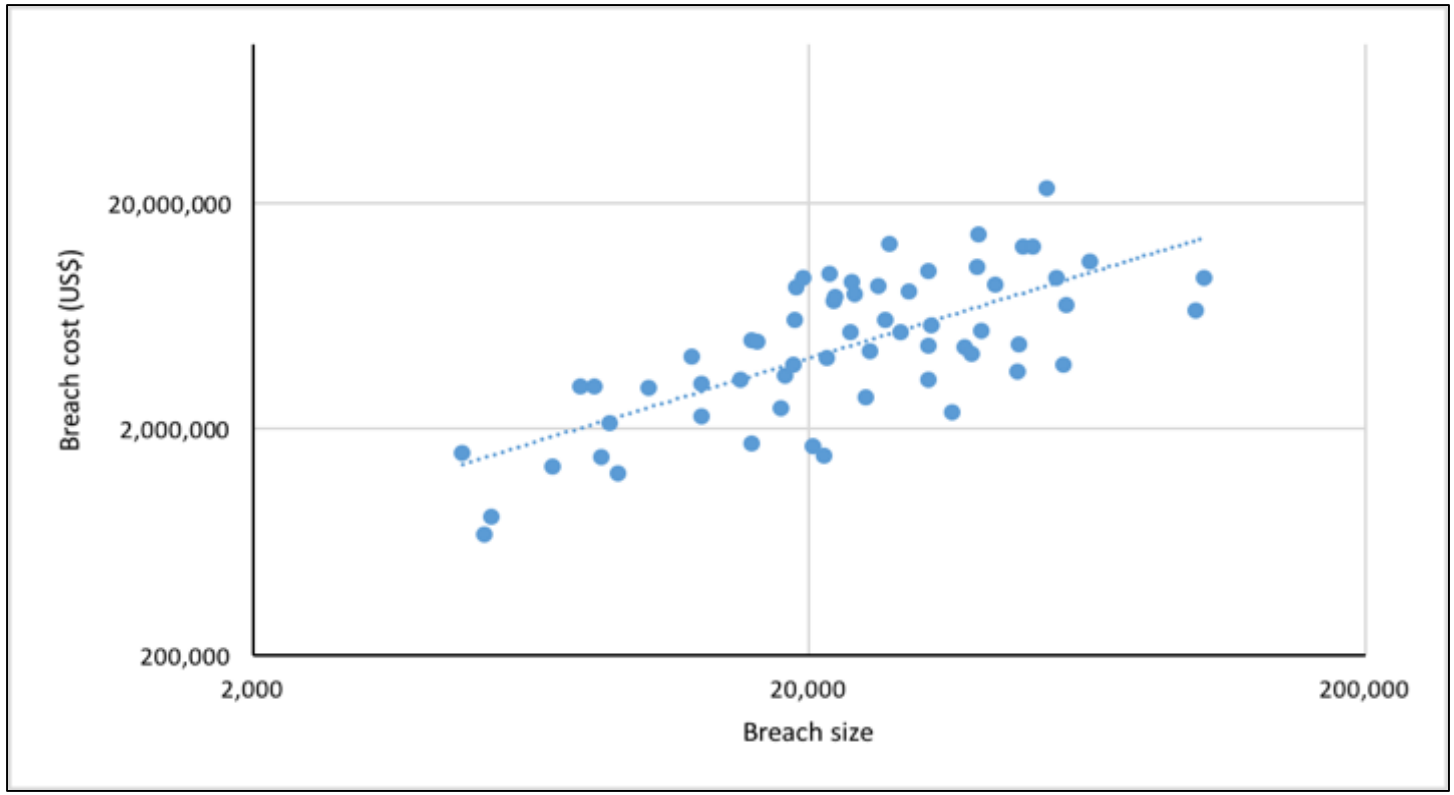
"Why Ponemon Institute's Cost of Data Breach Methodology Is Sound and Endures". Ponemon Institue. 2015.

The breach cost vs. breach size



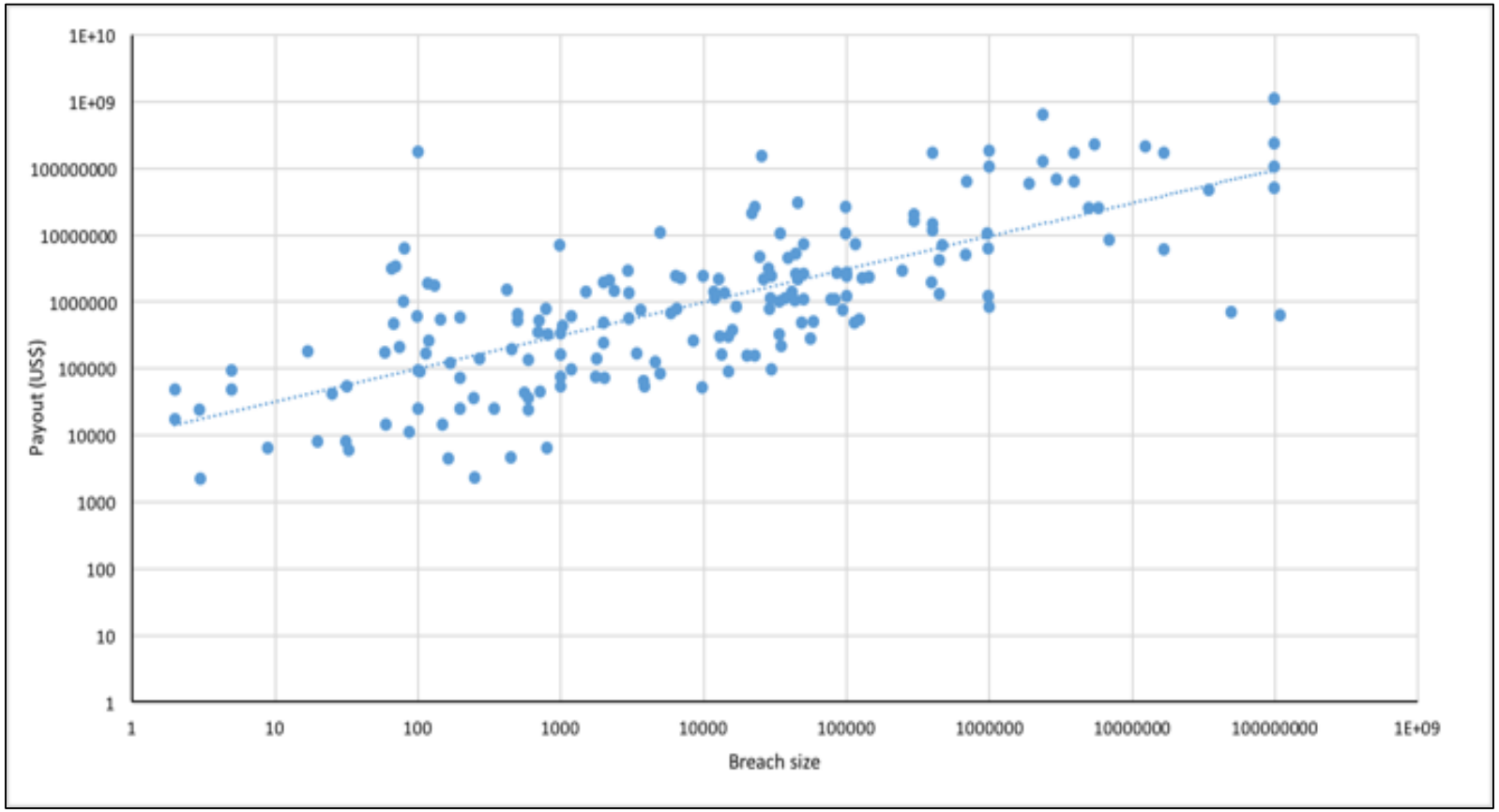
Ponemon 2013 data, the breach cost vs. breach size. Note log-log scate. (ranges from 5,000 to 100,000 records)

The breach cost vs. breach size



Ponemon 2014 data, the breach cost vs. breach size
(ranges from 4,700 to 103,000 records)

The breach cost vs. breach size



Verizon 2015 data, the claim amount vs. breach size
(ranges from single digits to 108 million records)

Colorado State University

The breach cost vs. breach size

- Our proposed model

$$\boldsymbol{Total\ breach\ cost} = a\ *\ size\ \hat{}\ b$$

for breach sizes bigger than or equal to 1000 records

- Nonlinearity caused by ***economy of scale,*** *thus b should be < 1.*
- Thus

$$\boldsymbol{Cost\ per\ record} = a\ *\ (size)\ \hat{}\ (b-1)$$

Annual Cost Models

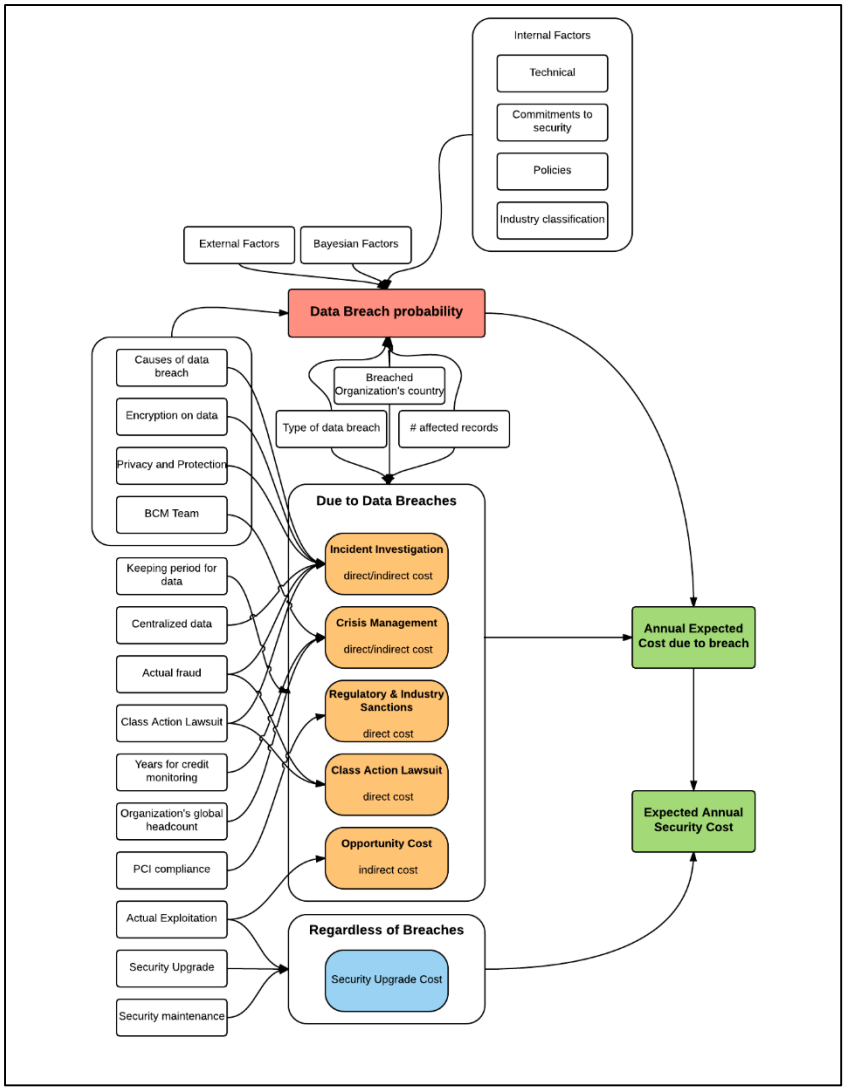
- ***Expected Annual Security Cost*** $=$
  Annual expected costs due to breaches $+$
  Costs regardless of any breaches
- ***Annual Expected Cost due to Breach*** $=$
  Σ Probability of a breach of data type i $\times$
  Total cost per breach for type i

Breach Cost/Payout Regression Models

Data sets	Size of breaches	Data points	Regression	
			Breach Cost Model	R^2
Ponemon 2013	5000 - 100,000	54	$y = 1924.2x^{0.7662}$	0.52
Ponemon 2014	4700 - 103,000	61	$y = 2439.9x^{0.7499}$	0.50
NetDiligence (Verizon report)	2 -108 million	183	$y = 10002x^{0.4971}$	0.54

Note: R^2 of 0.5 suggests moderate correlation. There are other factors that impact cost.

Overall risk evaluation model



Details in Abdullah Algarni's dissertation: Quantitative economics of security: software vulnerabilities and data breaches, CSU

Partial costs

- ***Investigation cost per record***

$$= [a.(size)^{b-1} \ for \ factors \ 4,5,6] * F_{breach\_cause} * F_{encryption} * F_{privacy}$$

- ***Crisis Management Cost per Record***

$$= [a.(size)^{(b-1)} \ for \ factor \ 11] * FBCM$$

- ***Sanctions cost per record***

$$= a.(size)^{b-1} \ for \ factor \ 14$$

- ***Class Action Lawsuit Cost per record***

$$= a.(size)^{b-1} \ for \ factor \ 15 \ and \ 16$$

- ***Opportunity cost: considered separately***

Conclusions

- Risk as Likelihood x Impact product
- Conditional components of Likelihood
- Vulnerability discovery and lifecycle
- CVSS as a risk measure: not completely validated yet
- Measuring impact: needs further research

References

1. O. H. Alhazmi, Y. K. Malaiya , I. Ray, "Measuring, Analyzing and Predicting Security Vulnerabilities in Software Systems," Computers and Security Journal, Volume 26, Issue 3, May 2007, Pages 219-228.
2. A. M. Algarni, Y. K. Malaiya,"Software Vulnerability Markets: Discoverers and Buyers," Int. Journal of Computer, Information Science and Engineering, Vol:8 No:3, 2014, pp. 71-81.
3. A. A. Younis, Y. K. Malaiya, and I. Ray, "Using Attack Surface Entry Points and Reachability Analysis to Assess the Risk of Software Vulnerability Exploitability", Proc. 2014 IEEE 15th International Symposium on High-Assurance Systems Engineering (HASE 2014), Miami, January, 2014, pp. 1-8.
4. A. M. Algarni and Y.K. Malaiya, "A Consolidated Approach for Estimation of Data Security Breach Costs", 2nd Int. Conf. on Information Management (ICIM), London, 2016, pp. 26-39
5. A. Younis, Y. Malaiya and I. Ray, "Evaluating CVSS Base Score Using Vulnerability Rewards Programs",Proc. 31th Int. Information Security and Privacy Conference, IFIP SEC, Ghent, Belgium, 2016, pp. 62-75.