Why Cyber Risk Measurement and Modeling Matters and How to Get it Done Securely

Daniel Weitzner
3Com Founders Senior Research Scientist
Massachusetts Institute of Technology Internet Policy Research Initiative

6.6550 -- May 2. 2024
Overview

A cybersecurity research quest
- Analyzing mandatory key escrow for law enforcement surveillance
- Critical Infrastructure resource allocation questions
  - 2014 report
  - Fed conference
- Models for how society deals with large-scale, dynamic risk
- Challenges of measure cyber security risk
- Early lessons
- Policy implications
- Technical implications
The Encryption & Surveillance Debate as a Test of Security Metrics
EU Encryption Policy History

<1992: Encryption regulated as a ‘munition’
1992: S.266 - plain text must be available when authorized by law
1993: Clipper Chip Proposed
1994: Blaze breaks Clipper (1995: Yung and Frankel show other vulns)
1994: Congress enacts CALEA, ‘information services’ and encryption exempted
1997; Nail in coffin -- Risks of Key Recovery
2010: CALEA2 proposed by FBI -- rejected by Obama White House
2013: Snowden → iOS & Android device encryption by default & https everywhere
==
2014: Comey ‘going dark again’
2015: Keys Under Doormats & Don’t Panic
2016: House of Representatives Crypto Report
2021: Apple proposed (then retracts) CSAM scanning for e2e encrypted messaging
Edward Snowden
Apple: “Unlike our competitors, Apple cannot bypass your passcode and therefore cannot access this data,” Apple said on its Web site. ‘So it’s not technically feasible for us to respond to government warrants for the extraction of this data from devices in their possession running iOS 8.’” (WaPo 9/18/2014)

Google: “The next generation of Google’s Android operating system, due for release next month, will encrypt data by default for the first time, the company said Thursday, raising yet another barrier to police gaining access to the troves of personal data typically kept on smartphones.”
Going Dark

“Those charged with protecting our people aren’t always able to access the evidence we need to prosecute crime and prevent terrorism even with lawful authority. We have the legal authority to intercept and access communications and information pursuant to court order, but we often lack the technical ability to do so.”

The issue is whether companies not currently subject to the Communications Assistance for Law Enforcement Act should be required to build lawful intercept capabilities for law enforcement. We aren’t seeking to expand our authority to intercept communications. We are struggling to keep up with changing technology and to maintain our ability to actually collect the communications we are authorized to intercept.

FBI Director James Comey (2014)

“The hope, perhaps, is that Silicon Valley, having engineered a problem, might just engineer a solution too.”

Apple vs FBI

Apple encryption debate after San Bernardino terrorist attack - IPRI contribution to policy conversation:

Keys Under Doormat paper

Findings of Keys Under Doormats paper

The deployment of key-recovery-based encryption infrastructures to meet law enforcement’s stated specifications will result in substantial sacrifices in security and greatly increased costs to the end user:

1. Loss of forward secrecy: required to keep keys around too long
2. Increased complexity: axiomatically bad for security
3. Centralized attack points

Building the secure computer-communication infrastructures necessary to provide adequate technological underpinnings demanded by these requirements would be enormously complex and is far beyond the experience and current competency of the field. Even if such infrastructures could be built, the risks and costs of such an operating environment may ultimately prove unacceptable.

Washington Post Editorial

How to resolve this? A police ‘back door” for all smartphones is undesirable — a back door can and will be exploited by bad guys, too. However, with all their wizardry, perhaps Apple and Google could invent a kind of secure golden key they would retain and use only when a court has approved a search warrant.

WaPo, October 3, 2014
Trend: Consensus shifts away from mandatory back doors

US Secretary of Defense Ash Carter: There will not be some simple, overall technical solution—a so-called ‘back door’ that does it all…. I’m not a believer in backdoors or a single technical approach. I don’t think that’s realistic.

UK GCHQ Director Robert Hannigan: The solution is not, of course, that encryption should be weakened, let alone banned. But neither is it true that nothing can be done without weakening encryption. I am not in favour of banning encryption just to avoid doubt. Nor am I asking for mandatory backdoors.

Emanuel Commission Vice-President Anders Ansip: “How will people trust the results of the election if they know that the government has a back door into the technology used to collect citizen’s votes?”

US House of Representatives Encryption Working Group: Cryptography experts and information security professionals believe that it is exceedingly difficult and impractical, if not impossible, to devise and implement a system that gives law enforcement exceptional access to encrypted data without also compromising security against hackers, industrial spies, and other malicious actors.
Alternative Approaches - UK, Australia, India

Policy Milestones - US
1992: S.266 - plain text must be available when authorized by law
1993: Clipper Chip Proposed
1994: CALEA, ‘information services’ and encryption exempted
1997: Nail in coffin -- Risks of Key Recovery

2010: CALEA2 proposed by FBI -- rejected by Obama WH
2013: Snowden disclosures
2014: iOS & Android device encryption by default & https everywhere
2014: Comey ‘going dark again’
2015: Keys Under Doormats & Don’t Panic
2016: House of Reps Crypto Report
2018: AG Barr

2010: India demands Blackberry provide exceptional access


2018: Australia - Assistance and Access Bill

2020: India: Proposed filtering and decryption requirements on Internet platforms
Next-wave encryption surveillance policy: Legislators move policy debate to opaque regulatory process and ‘experts’

UK Investigative Powers Act of 2016

- Comprehensive surveillance law reform
- **Technical Capacity Notices** authorized against all ‘service providers’
  - Must be ‘technically reasonable’
  - Evaluated by Information Commission and technical advisory board
- Technical requirements and evaluations may be kept secret
- Has power been exercised?

Australian Assistance and Access Bill (2018)

- “Technical Capacity Notices”
- Systemic Vulnerabilities disallowed
- Criminal offense to disclose Technical Capacity Notices

MIT IPRI Experts Letter on risks to security transparency features:

1. Certificate Transparency
2. Message Key Transparency
3. Binary Transparency
What does the ‘expert’ debate look like?

“[P]utting aside the more controversial debate about data in motion ... and focusing on a conversation about data at rest ... allowed us to find a more pragmatic way to address the concerns of both privacy advocates and law enforcement. This was an important starting point, and while we did not conclude with an agreed upon proposal, we were able to make progress. Embracing this approach could help move this entrenched debate in a more constructive direction.”

Susan Landau, Denis McDonough, Breaking the encryption impasse, The Hill (1/16/2020)

“Proposals should be tested multiple times—including at strategic levels (for example, do they establish high-level principles and requirements to weed out incomplete or unfeasible proposals?) and at technical levels (for example, what are the technical risks of the specific implementation?).”

Questions we don’t yet know how to answer

Technical Questions

● What is the right measure of ‘technically feasibility’?
● How do we know when a EA requirement creates a vulnerability that is ‘systemic’?
● How can we measure relative security properties of systems with and without EA?
● Can security vulnerabilities be detected and evaluated in secret?

Policy Questions

● How do we assess the relative risks of (a) Exceptional Access system that could open up new vulnerabilities vs (b) limiting law enforcement access to investigative material?
● Do all TCNs have to be secret?
● What is the effect of secrecy on user trust and technical security properties?
Not Even the Most Well-Resourced Firms Know How to Measure The Cybersecurity Risk
Critical Infrastructure Security reveals knowledge gaps

SECOND CHALLENGE
Measure cyber risk and infrastructure fragility.

Finding:

Quantifying risk in either absolute or relative terms is a difficult challenge that impedes cybersecurity investment in all sectors examined except certain financial institutions. The asserted inability to measure the rate of return on cybersecurity investment is a closely related problem that affects overall investment levels and makes it difficult to target investment. Absent assurances of confidentiality, candid participation by the private sector will not occur. However, the public should be informed of the general state of security of critical infrastructure.
Tom Barkin, the President of the Federal Reserve Bank of Richmond, called for the industry, governments, and academia to work together to develop taxonomies and metrics to bring cyber risk closer to the models we already have in the financial service sector for managing operational and credit risk.
Current Cybersecurity Practices Are Walking Down Blind Alleys
Bank of America: No budget constraint for cyber

Bank of America Corp. CEO Brian Moynihan said [...] it was the first time in 20 years of corporate budgeting he had overseen a business unit [cybersecurity] with no budget. Moynihan said the only place in the company that didn't have a budget constraint was cybersecurity.

Core need: How to quantify security ROI?

2019 spend ~$700 million on cyber defense
Ransomware Attack Hits Martha’s Vineyard Ferry Service
Cybersecurity is not the first complex risk in society...

The Philadelphia Contributorship Fire Code: (technical design standards)

- must display a firemark
- must have a trap door to the roof to fight chimney and roof fires
- no trees in front of the house
- higher rates for high-risk businesses (breweries, apothecary)
- no adjourning bakehouses

Volunteer fire companies more likely to fight fires of Contributorship members
US Automobile Safety

United States Automobile Deaths

- Deaths per 100M VMT
- Fatalities per 100k population

Year

Deaths

US Auto Safety Regulatory Structure

- Government safety regulations
- Industry & government testing
- Insurance Companies
- Automobile Manufacturers

Evolving Safety Standards
### Vehicles with the lowest collision losses

<table>
<thead>
<tr>
<th>Vehicle</th>
<th>Size and class</th>
<th>Losses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jeep Wrangler 2dr 4WD</td>
<td>Midsize SUV</td>
<td>-64%</td>
</tr>
<tr>
<td>Chevrolet Express 2500 series</td>
<td>Very large van</td>
<td>-59%</td>
</tr>
<tr>
<td>Ford F-250 4WD</td>
<td>Very large pickup</td>
<td>-56%</td>
</tr>
<tr>
<td>Chevrolet Express 3500 series</td>
<td>Very large van</td>
<td>-54%</td>
</tr>
<tr>
<td>Ford F-250 SuperCab 4WD</td>
<td>Very large pickup</td>
<td>-49%</td>
</tr>
<tr>
<td>Ram 2500 4WD</td>
<td>Very large pickup</td>
<td>-47%</td>
</tr>
<tr>
<td>Jeep Wrangler 2dr SWB 4WD</td>
<td>Small SUV</td>
<td>-46%</td>
</tr>
<tr>
<td>Ford F-150 4WD</td>
<td>Large pickup</td>
<td>-45%</td>
</tr>
<tr>
<td>Fiat 500 convertible</td>
<td>Mini two-door car</td>
<td>-44%</td>
</tr>
<tr>
<td>Fiat 500 electric</td>
<td>Mini two-door car</td>
<td>-43%</td>
</tr>
</tbody>
</table>

### Vehicles with the highest collision losses

<table>
<thead>
<tr>
<th>Vehicle</th>
<th>Size and class</th>
<th>Losses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bentley Flying Spur 4dr 4WD</td>
<td>Very large luxury car</td>
<td>628%</td>
</tr>
<tr>
<td>Ferrari 488 GTB 2dr</td>
<td>Midsize sports car</td>
<td>464%</td>
</tr>
<tr>
<td>Bentley Bentayga 4dr 4WD</td>
<td>Large luxury SUV</td>
<td>451%</td>
</tr>
<tr>
<td>Ferrari 488 GTS convertible</td>
<td>Midsize sports car</td>
<td>375%</td>
</tr>
<tr>
<td>Alfa Romeo Giulia Quadrifoglio 4dr</td>
<td>Midsize luxury car</td>
<td>358%</td>
</tr>
<tr>
<td>Audi R8 2dr 4WD</td>
<td>Large sports car</td>
<td>335%</td>
</tr>
<tr>
<td>Audi S8 4dr 4WD</td>
<td>Very large luxury car</td>
<td>233%</td>
</tr>
<tr>
<td>Rolls Royce Dawn convertible</td>
<td>Very large luxury car</td>
<td>315%</td>
</tr>
<tr>
<td>Maserati Quattroporte 4dr</td>
<td>Very large luxury car</td>
<td>308%</td>
</tr>
<tr>
<td>Nissan GT-R 2dr 4WD</td>
<td>Midsize sports car</td>
<td>285%</td>
</tr>
</tbody>
</table>

Insurance losses by make and model 2011-2013 (IIHS/HLDI)
Current legal, policy, institutional approaches
Case study in data-free policy making: UK Information Commissioner fine against British Airways

- 496,635 data subjects’ info breached
- original penalty £183 Million (July 2019)
  - reduced to £30 Million (October 2020)
  - further reduced to £24 Million factoring in loss of BA’s reputation
  - even further reduced to because of COVID hardship
  - final result £20 Million Penalty

Metric: \[ SE = DPC \times EI + CB \]
- SE = Overall Severity
- DPC = Data Processing Context
- EI = Ease of Identification
- CB = Circumstances of Breach

<table>
<thead>
<tr>
<th>Severity of a data breach</th>
<th>SE</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>SE &lt; 2</td>
<td>Individuals either will not be affected or may encounter a few inconveniences, which they will overcome without any problem (time spent re-entering information, annoyances, irritations, etc.).</td>
</tr>
<tr>
<td>Medium</td>
<td>2 \leq SE &lt; 3</td>
<td>Individuals may encounter significant inconveniences, which they will be able to overcome despite a few difficulties (extra costs, denial of access to business services, fear, lack of understanding, stress, minor physical ailments, etc.).</td>
</tr>
<tr>
<td>High</td>
<td>3 \leq SE &lt; 4</td>
<td>Individuals may encounter significant consequences, which they should be able to overcome albeit with serious difficulties (misappropriation of funds, blacklisting by banks, property damage, loss of employment, subpoena, worsening of health, etc.).</td>
</tr>
<tr>
<td>Very High</td>
<td>4 \leq SE</td>
<td>Individuals may encounter significant, or even irreversible, consequences, which they may not overcome (financial distress such as substantial debt or inability to work, long-term psychological or physical ailments, death, etc.).</td>
</tr>
</tbody>
</table>

ENISA “Recommendations for a methodology of the assessment of severity of personal data breaches” Working Document, v1.0, December 2013
Modeling Security Risk With Secure Multi-party Computation/FHE Platforms
SCRAM Research Project Motivations

Cybersecurity strategy & policy questions we cannot answer today:

1. How should enterprises prioritize cybersecurity investments?
2. What insurance underwriting models will yield efficient results?
3. How can we evaluate the efficiency of cybersecurity regulatory requirements?

Loss event magnitude

MIT Ransomware Readiness Index (RRI)

Cyber Risk Measurement for Municipalities
Security control maturity, by control, percentage of municipalities

1a. Deploy multi-factor authentication across the enterprise
1b. Deploy an endpoint detection and response (EDR) system / host-based IPS agent
2a. Hunt for malicious activity
2b. Encrypt data in transit
3a. Encrypt data at rest
3b. Remove barriers to sharing threat intelligence
4a. Receive external threat intelligence
4b. Evaluate employee skills
5a. Deliver regular training
5b. Protect backups
6a. Perform regular backups of systems
6b. Test backup data
6c. Store backups in offline location
6d. Implement a centralized backup management system
7a. Apply patches using a risk-based approach
7b. Establish an external penetration testing program
7c. Maintain your incident response plan
8a. Adopt network segmentation to ensure isolation of critical systems in an attack
8b. Perform red team exercises
8c. Adopt network segmentation to ensure isolation of critical systems in an attack
8d. Test your incident response plan
9a. Codify an incident response plan
9b. Maintain your incident response plan
9c. Establish an external penetration testing program
10a. Adopt network segmentation to ensure isolation of critical systems in an attack
10b. Test your incident response plan
Losses from failures, by control, total, USD

1a. Deploy multi-factor authentication across the enterprise
2a. Deploy an endpoint detection and response (EDR) system / host-based IPS agent
   2b. Hunt for malicious activity
   3a. Encrypt data in transit
   3b. Encrypt data at rest
4a. Remove barriers to sharing threat intelligence
   4b. Receive external threat intelligence
5a. Evaluate employee skills
   5b. Deliver regular training
   6a. Perform regular backups of systems
   6b. Test backup data
   6c. Protect backups
   6d. Store backups in offline location
7a. Deploy updates and patches in a timely manner
   7b. Implement a centralized patch management system
   7c. Apply patches using a risk-based approach
8a. Codify an incident response plan
   8b. Test your incident response plan
   8c. Maintain your incident response plan
9a. Establish an external penetration testing program
   9b. Perform red team exercises
10a. Adopt network segmentation to ensure isolation of critical systems in an attack

$0  $20,000  $40,000  $60,000  $80,000  $100,000  $120,000  $140,000
Frequency of control failure, by control, total

1a. Deploy multi-factor authentication across the enterprise
2a. Deploy an endpoint detection and response (EDR) system / host-based IPS agent
2b. Hunt for malicious activity
3a. Encrypt data in transit
3b. Encrypt data at rest
4a. Remove barriers to sharing threat intelligence
4b. Receive external threat intelligence
5a. Evaluate employee skills
5b. Deliver regular training
6a. Perform regular backups of systems
6b. Test backup data
6c. Protect backups
6d. Store backups in offline location
7a. Deploy updates and patches in a timely manner
7b. Implement a centralized patch management system
7c. Apply patches using a risk-based approach
8a. Codify an incident response plan
8b. Test your incident response plan
8c. Maintain your incident response plan
9a. Establish an external penetration testing program
9b. Perform red team exercises
10a. Adopt network segmentation to ensure isolation of critical systems in an attack
Evidence-based Policy Guidance

Red: Low maturity, low adoption, and high losses. Should be a primary areas of focus

Yellow: Middle ground. Good in some, worse in others. Still potential for large losses

Green: Most mature areas in list, but still room for improvement.

*refers to information sharing and intelligence gathering

Analysis by:
- Taylor Reynolds (MIT, treyn@mit.edu)
- Chelsea Conard (MIT)

Large score: Out of 100 (higher = better)
MIT data: n = 84
Insurance survey data (b): n = 215
Details under score
Mat = Maturity rating, self rated by organizations (0-100)
Adopt = Percent adoption, normalized. (0-100)
a = MIT survey using SCRAM, status June 2022
b = Insurance survey, 2023
a & b are not directly compatible, use different methodologies
Loss = Financial losses from MIT data, categorized, normalized (20-80)
Calculating an organization’s expected risk and optimally efficient investment level
**PLG = R**

If the organization is prepared to cover a loss with a magnitude of \(\frac{R}{P}\), then its actions would be considered reasonable.

- **Probability** \(P\)
- **Loss** \(L\)
- **Gap Index** (Defense) \(G\)

\[
P \times L \times G = R
\]

- **R/P** incident size (forecasted)
- **Duty of security threshold**
- **Reasonableness?**

Duties of security threshold

- Adjusted by firm’s actual defenses (1 = no difference)
- Firm Specific (annual R)
- Firm Specific (incident)
Example 1: Muni with **average** security

<table>
<thead>
<tr>
<th></th>
<th>Calculated for the sector (annual P)</th>
<th>Calculated for the sector (average)</th>
<th>Baseline for the sector (average)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>Probability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>Loss</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>Gap Index (Defense)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>Risk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R/P</td>
<td>Incident size (forecasted)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
P \times L \times G = R \rightarrow R/P
\]

0.016 * $157,688 * 1 = $2,523 \rightarrow $157,688

**Reasonableness?**
If this municipality is prepared to cover a loss of R/P ($157,688), then its actions would be considered reasonable.
Example 2: Muni with 30% lower security

Calculated for the sector (annual $P$) * Calculated for the sector (average) * Baseline for the sector (average) = R/P Incident size (forecasted)

$P$ * $L$ * $G$ = $R$ → $R/P$

Probability * Loss * Gap Index (Defense) = Risk → Incident size (forecasted)

0.016 * $157,688 * 4.75 = $11,980 → $748,729

Reasonableness?
If this municipality is prepared to cover a loss of $R/P$ ($748,729), then its actions would be considered reasonable.
Example 3: Muni with 10% higher security

\[ P \times L \times G = R \]

\[ \frac{R}{P} \]

- **Probability** (Defense)
- **Loss**
- **Gap Index (Defense)**
- **Risk**
- **R/P Incident size (forecasted)**

\[ P = 0.016 \]
\[ L = \$157,688 \]
\[ G = 0.59 \]
\[ R = \$1,493 \]
\[ \frac{R}{P} = \$93,315 \]

**Reasonableness?**
If this municipality is prepared to cover a loss of R/P ($93,315), then its actions would be considered reasonable.

- Adjusted by firm’s actual defenses (1 = no difference)
- Firm Specific (annual R)
- Firm Specific (incident)
Incident sizes grow quickly with poor security

Predicted municipal security incident size, by net weighted security control deviation from group average, USD

<table>
<thead>
<tr>
<th>Deviation</th>
<th>Size (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-35%</td>
<td>$971,340</td>
</tr>
<tr>
<td>-30%</td>
<td>$748,729</td>
</tr>
<tr>
<td>-25%</td>
<td>$577,135</td>
</tr>
<tr>
<td>-20%</td>
<td>$444,868</td>
</tr>
<tr>
<td>-15%</td>
<td>$342,913</td>
</tr>
<tr>
<td>-10%</td>
<td>$264,324</td>
</tr>
<tr>
<td>-5%</td>
<td>$203,747</td>
</tr>
<tr>
<td>0%</td>
<td>$157,052</td>
</tr>
<tr>
<td>5%</td>
<td>$121,059</td>
</tr>
<tr>
<td>10%</td>
<td>$93,315</td>
</tr>
<tr>
<td>15%</td>
<td>$71,929</td>
</tr>
<tr>
<td>20%</td>
<td>$55,444</td>
</tr>
<tr>
<td>25%</td>
<td>$42,738</td>
</tr>
<tr>
<td>30%</td>
<td>$32,943</td>
</tr>
<tr>
<td>35%</td>
<td>$25,393</td>
</tr>
</tbody>
</table>
Empirical risk assessments for software libraries
Empirical analysis of software vulnerabilities indicating reasonable vs unreasonable behavior

Reasonable choice of software components: different SSL libraries have different failure rates.

Memory-safe languages produce uniformly more reliable security libraries

Figure 5: Vulnerability Comparison Post-Heartbleed—Vulnerabilities discovered in OpenSSL after the initial releases of LibreSSL and BoringSSL with a comparison of how many of those vulnerabilities also affected LibreSSL and BoringSSL.

Figure 1: Vulnerability Types — Taxonomy of vulnerabilities by root cause across six main categories within cryptographic software. Table 8 contains descriptions of the taxonomy used and examples of vulnerabilities from each category.

Federal Policy Proposal to Incentivize Efficient Investment Behavior
Proposal – A ‘reasonableness’ standard that promotes efficient security investment

Proposal: Establish a safe harbor from liability for firms whose cybersecurity practices are set according to a validated, empirical methodology that identifies on a dynamic basis efficient security strategies based on known control failures, insecure design, and observed losses from firms with similar risk profiles.

Rule: If an organization has allocated sufficient resources to address the amount of their expected risk and forecasted incident size based on a baseline of their peers and adjusted for their own security posture, then they should pass the reasonableness test/have met their duty.

Rationale: In the face of changing threats and vulnerabilities, reasonableness must be dynamically defined. Efficient allocation of cybersecurity resources can only be accomplished based on studying actual losses and control failures as revealed in actual sector-wide and cross-sector data.

Advantage over current approaches: Most cybersecurity frameworks depend on some notion of reasonableness, but ongoing dispute on how reasonableness should be defined:

- General reliance on industry best practices: adaptable to new threats but lacking certainly for firms
- Statutory or regulatory standards: likely too brittle
- Industry standards: predictability for firms but rarely empirically validated
- PL > B: hard to determine in dynamic and multi-party threat environment
CONCLUSION
Advertisement

6.4590 Foundations of Internet Policy

https://internetpolicy.mit.edu/6.4590/
References

• T. Reynolds, D. Weitzner, Mind the Gap: Securely modeling cyber risk based on security deviations from a peer group. (Under review, November 2023, pre-print available on request)

• A. Baral, T. Reynolds, L. Susskind, D. Weitzner, Cyber risk modeling for public policy and enterprise decision-making: A case study in municipal cyber risk measurement (Cybersecurity Law and Policy Scholars Conference, Sept 2023)

