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

Internet Policy Research Initiative

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

Silicon Valley Response to Snowden and USG Intrusion

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.”

United States Attorney General Loretta Lynch (2015)

Apple vs FBI

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

Keys Under Doormat paper

"All the News That's Fit to Print" **The New York Times**

VOL. CLXIV . . . No. 56,921 NEW YORK, WEDNESDAY, JULY 8, 2015

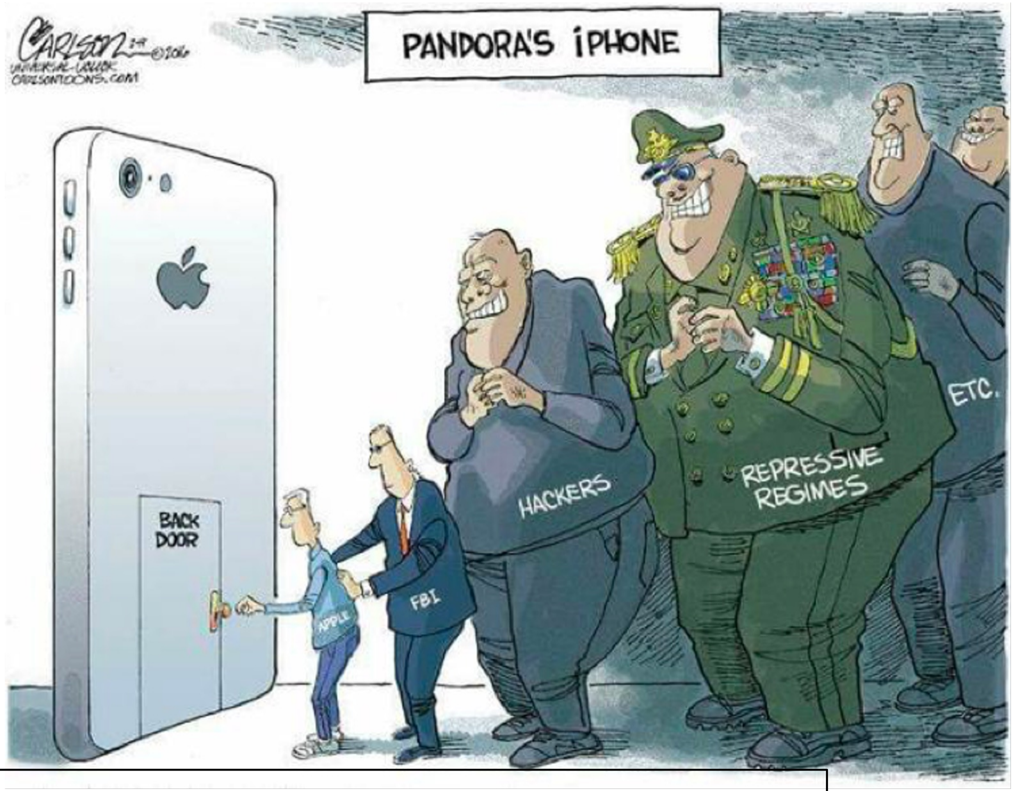
Security Experts Oppose Government Access to Encrypted Communication

By NICOLE PERLROTH JULY 7, 2015

SAN FRANCISCO — An elite group of security technologists has concluded that the American and British governments cannot demand special access to encrypted communications without putting the world's most confidential data and critical infrastructure in danger.

The Washington Post

Weitzner: Encryption solution in wake of Paris should come from Washington not Silicon Valley



Abelson, Rivest, Schiller, Specter, Weitzner, et al. "Keys under doormats: mandating insecurity by requiring government access to all data and communications." Journal of Cybersecurity 1.1 (2015): 69-79.




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.

Abelson, Rivest, Schiller, Specter, Weitzner, et al. "Keys under doormats: mandating insecurity by requiring government access to all data and communications." Journal of Cybersecurity 1.1 (2015): 69-79.

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

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.*

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.

European 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: Blaze breaks Clipper (1995: Yung and Frankel
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 discuslosures
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

2016: UK Investigative Powers Act: "Snoopers Charter"

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 disclosure 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?).”

-Carnegie Encryption Working Group, “Moving the Encryption Policy Conversation Forward.” (Sept 2019)

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

MIT CENTER FOR INTERNATIONAL STUDIES | Internet Policy Research Initiative

KEEPING AMERICA SAFE:
TOWARD MORE SECURE NETWORKS FOR CRITICAL SECTORS

Cambridge, Massachusetts
March 2017

Joel Brenner
brennerj@mit.edu

Report on a Series of MIT Workshops, 2015-2016

With Recommendations for the New Administration

MIT Center for International Studies
https://cis.mit.edu

MIT Internet Policy Research Initiative
https://internetpolicy.mit.edu

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.

MIT-Federal Reserve Collaboration Seeks Better Cyber Risk Metrics

Measuring Cyber Risk in the Financial Services Sector: Conference Summary (Executive Summary)

Sponsored by
The Board of Governors of the Federal Reserve System
The Federal Reserve Bank of Richmond
The Massachusetts Institute of Technology – Internet Policy Research Initiative

28 March 2024

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.

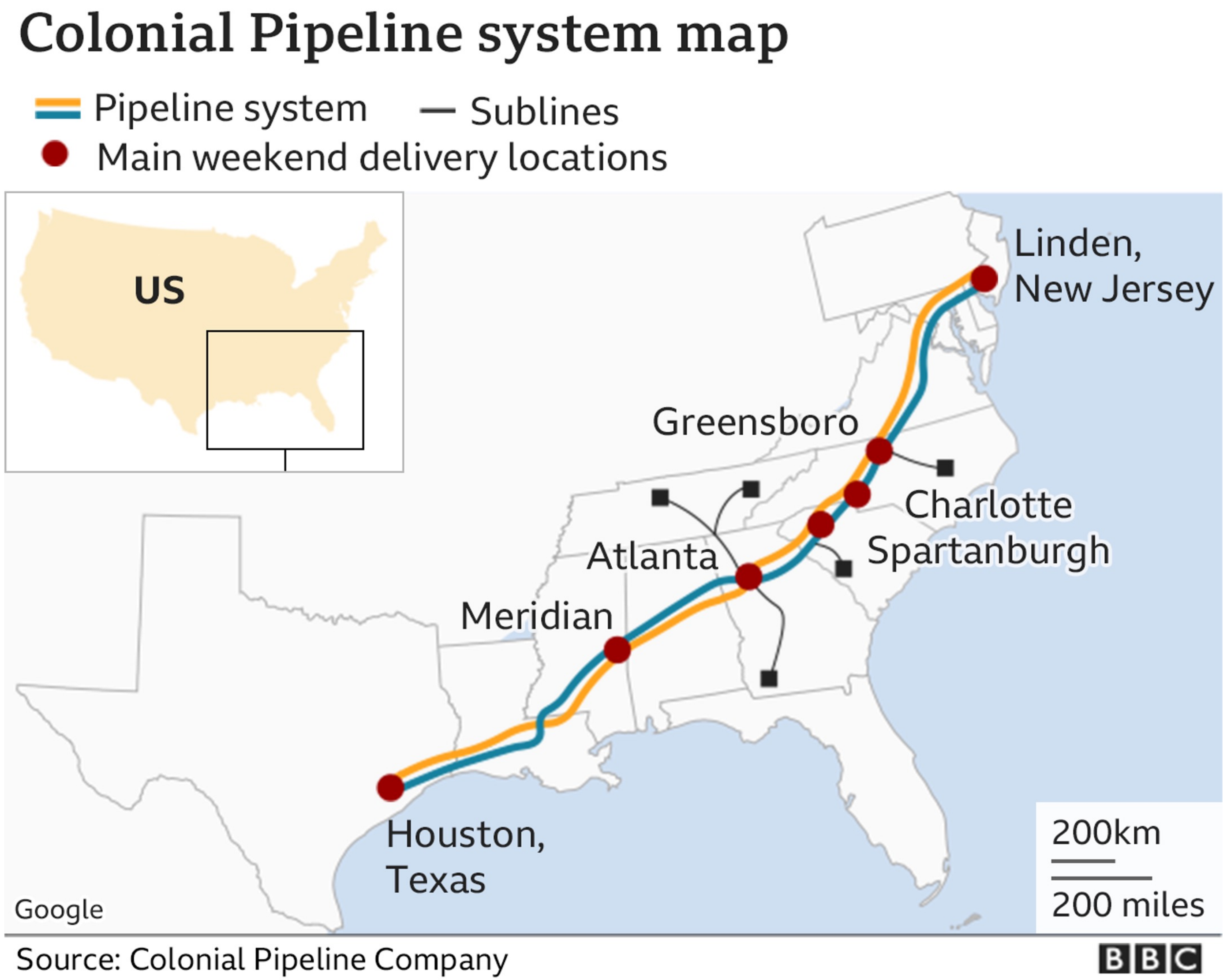
Core need: How to quantify security ROI?

2019 spend ~$700 million on cyber defense

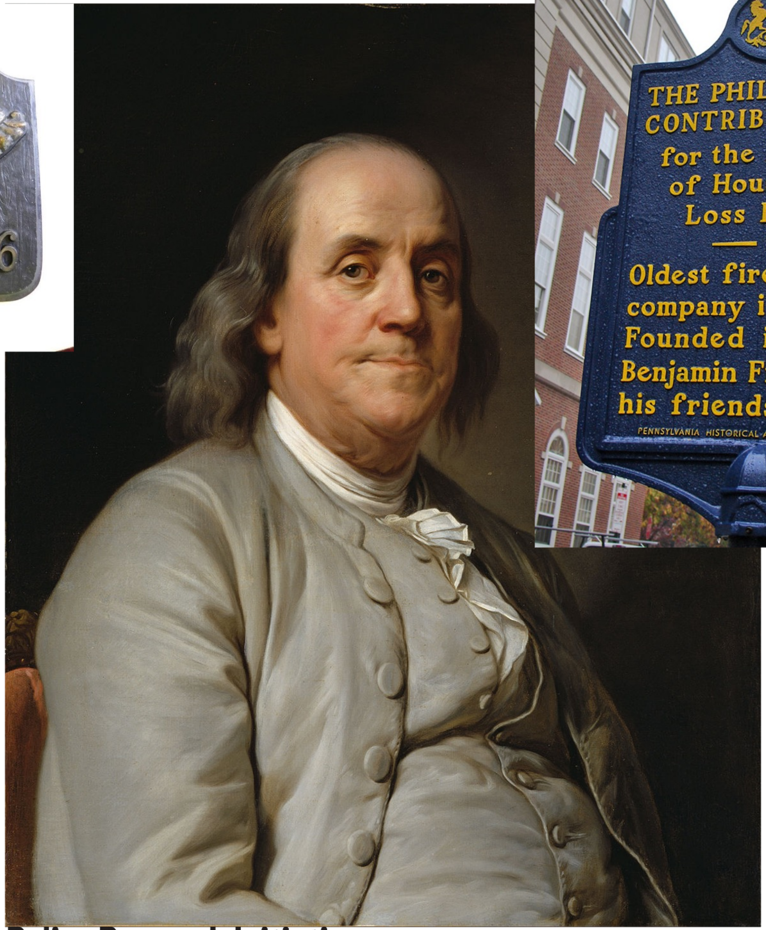
U.S.

Ransomware Attack Hits Martha's Vineyard Ferry Service


COLONIAL PIPELINE CO
1473


Colonial Pipeline system map
Pipeline system
Sublines
Main weekend delivery locations
US
Linden, New Jersey
Greensboro
Charlotte
Spartanburgh
Atlanta
Meridian
Houston, Texas
200km
200 miles
Google
Source: Colonial Pipeline Company
BBC

Cybersecurity is not the first complex risk in society…

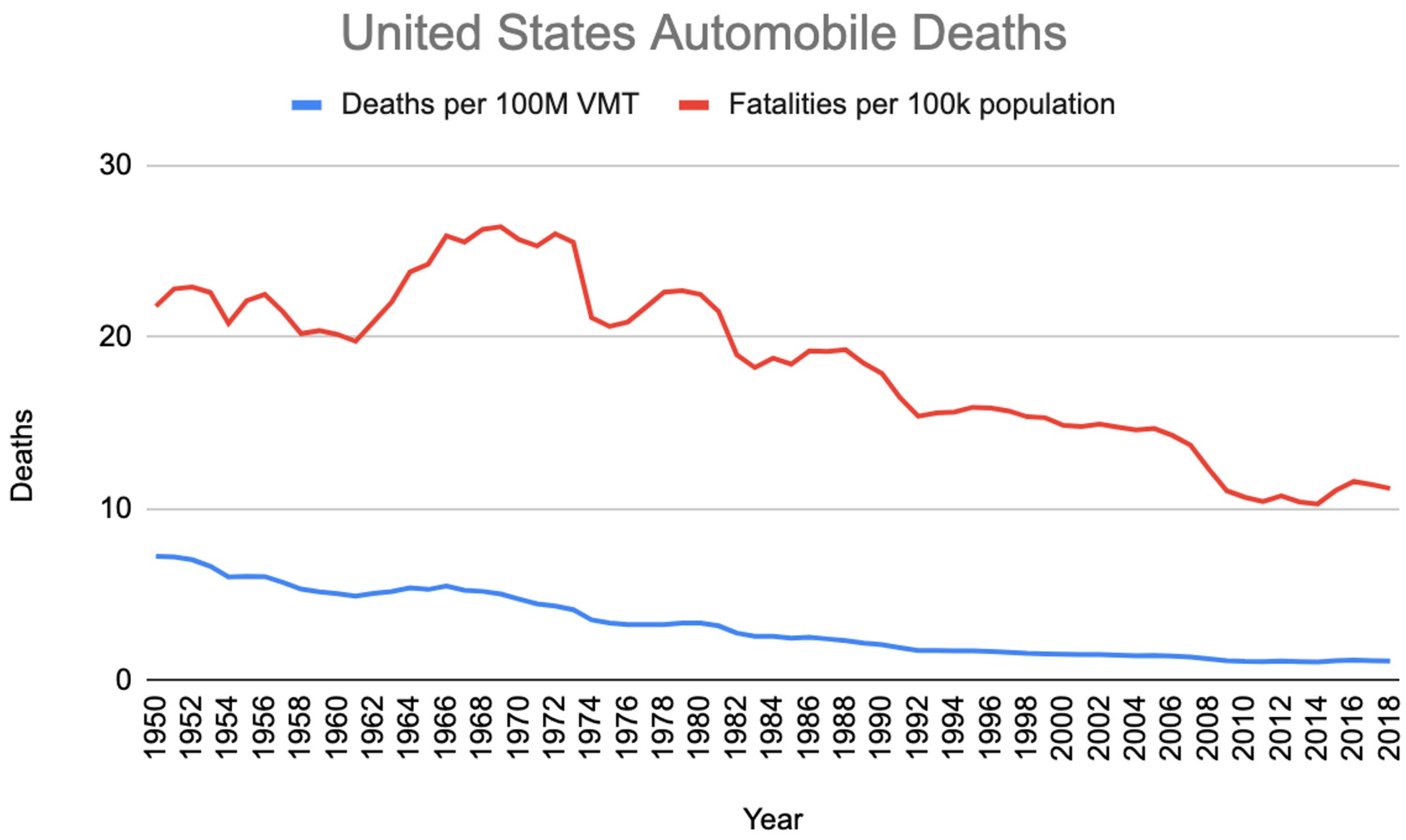


The Philadelphia Contributorship Fire Code:
(technical design standards)

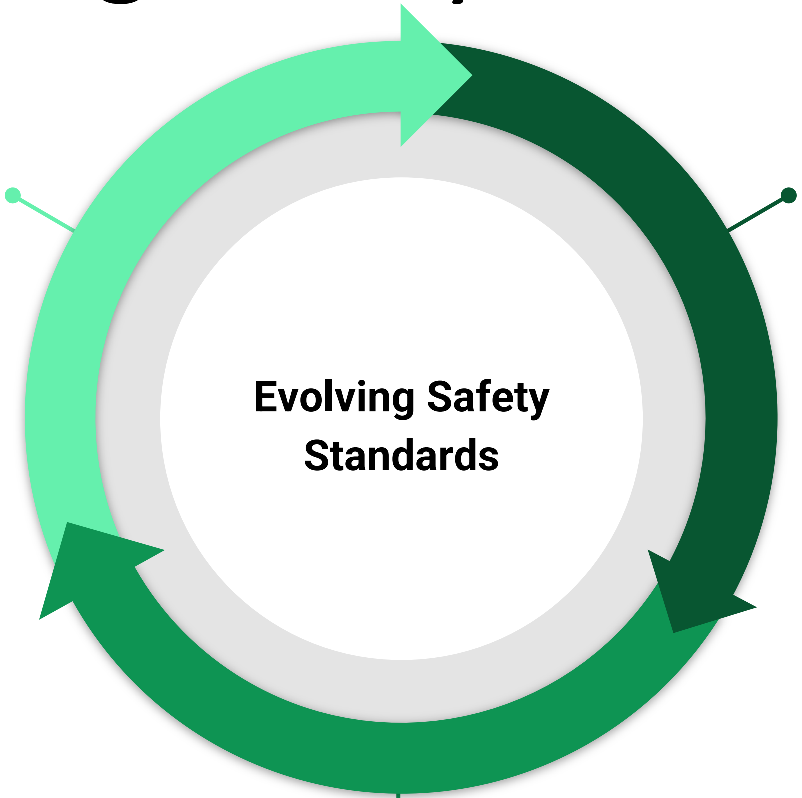
- must display a firemark
- must have a trap door to the roof to fight chimney and roof fires
- no trees in front of the house
- higher rates for high-risk businesses (breweries, apothecary)
- no adjourning bakehouses

Volunteer fire companies more likely to fight fires of Contributorship members

US Automobile Safety



US Auto Safety Regulatory Structure



Risk Pricing:Tracking insurance losses by vehicle type

Vehicles with the lowest collision losses

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

Vehicles with the highest collision losses

Vehicle	Size and class	Losses
Bentley Flying Spur 4dr 4WD	Very large luxury car	628%
Ferrari 488 GTB 2dr	Midsize sports car	464%
Bentley Bentayga 4dr 4WD	Large luxury SUV	451%
Ferrari 488 GTS convertible	Midsize sports car	375%
Alfa Romeo Giulia Quadrifoglio 4dr	Midsize luxury car	358%
Audi R8 2dr 4WD	Large sports car	336%
Audi S8 4dr 4WD	Very large luxury car	333%
Rolls Royce Dawn convertible	Very large luxury car	315%
Maserati Quattroporte 4dr	Very large luxury car	308%
Nissan GT-R 2dr 4WD	Midsize sports car	285%

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

ENISA "Recommendations for a methodology of the assessment of severity of personal data breaches" Working Document, v1.0, December 2013

Metric: SE = DPC x EI + CB

- SE = Overall Severity
- DPC = Data Processing Context
- EI = Ease of Identification
- CB = Circumstances of Breach

Severity of a data breach		
SE < 2	Low	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.).
2 ≤ SE < 3	Medium	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.).
3 ≤ SE< 4	High	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.).
4 ≤ SE	Very High	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.).

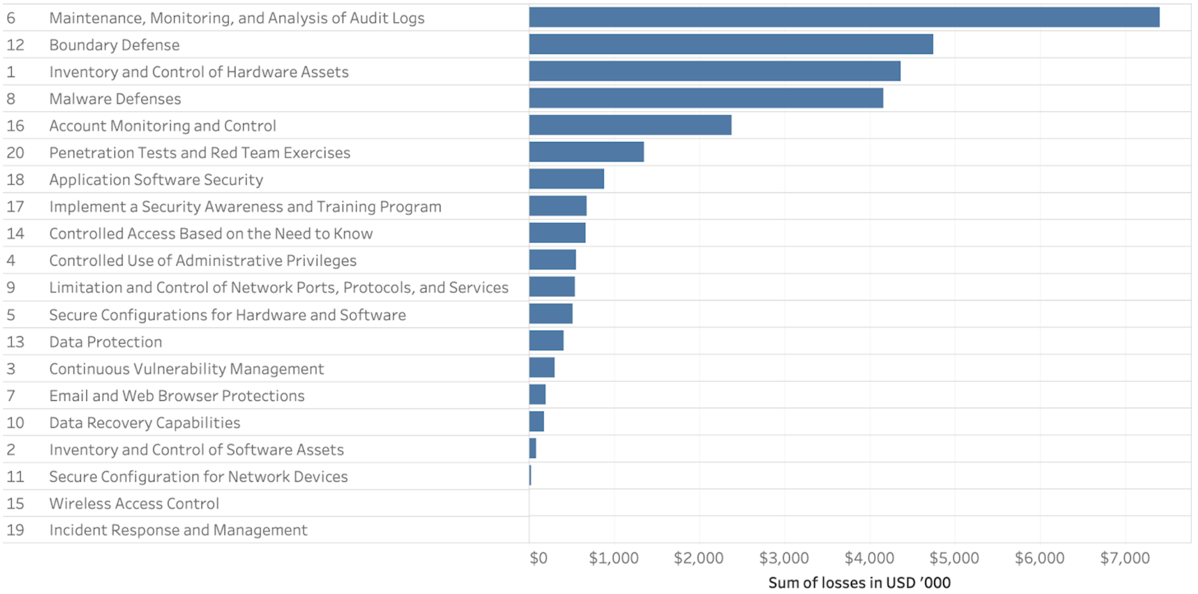
Modeling Security Risk With Secure Multi-party Computation/FHE Platforms

SCRAM Research Project Motivations

Cybersecurity strategy & policy questions we cannot answer today:

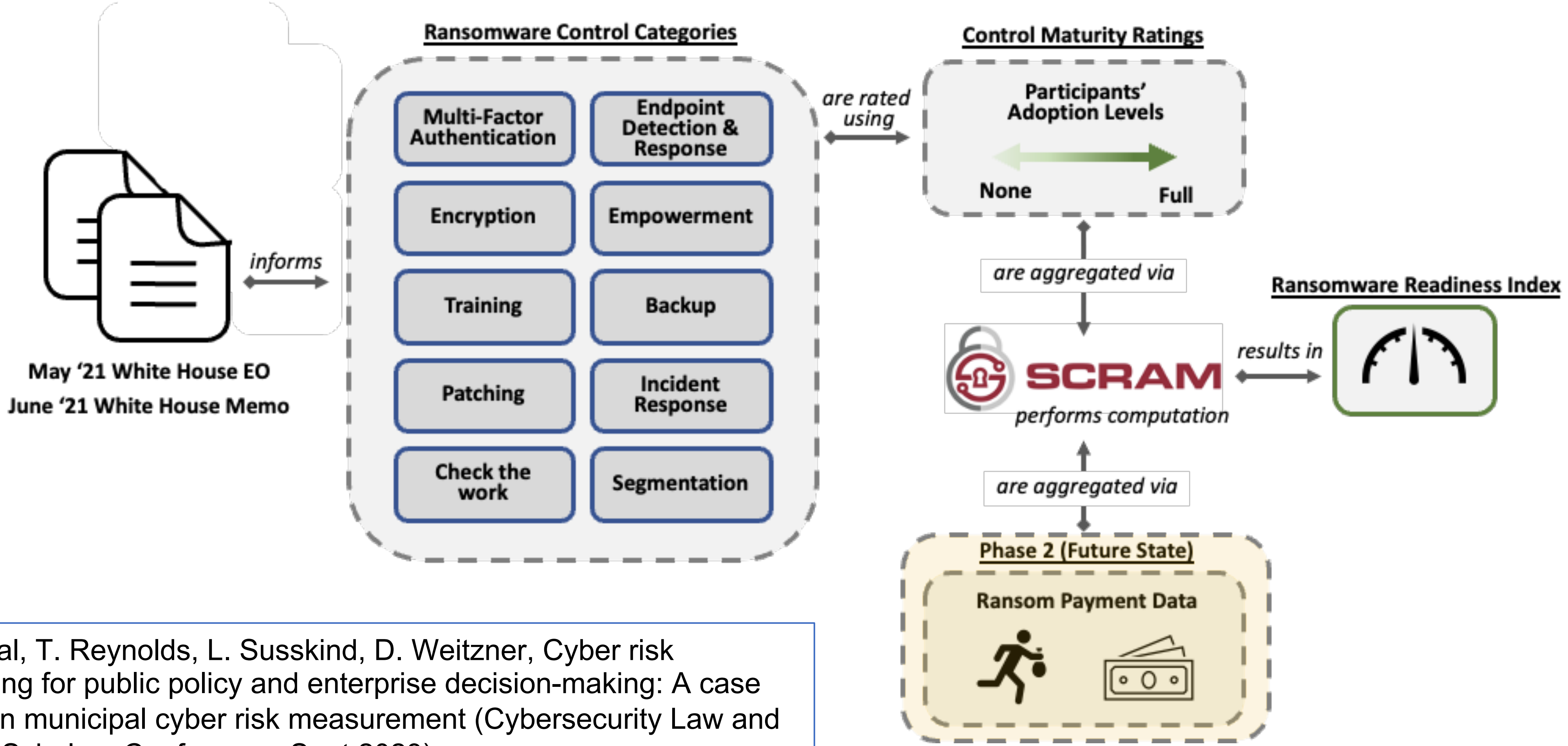
1. How should enterprises prioritize cybersecurity investments?
2. What insurance underwriting models will yield efficient results?
3. How can we evaluate the efficiency of cybersecurity regulatory requirements?

Loss event magnitude



de Castro, L., Lo, A. W., Reynolds, T., Susan, F., Vaikuntanathan, V., Weitzner, D., & Zhang, N. (2020). **SCRAM: A Platform for Securely Measuring Cyber Risk . Harvard Data Science Review**. https://doi.org/10.1162/99608f92.b4bb506a

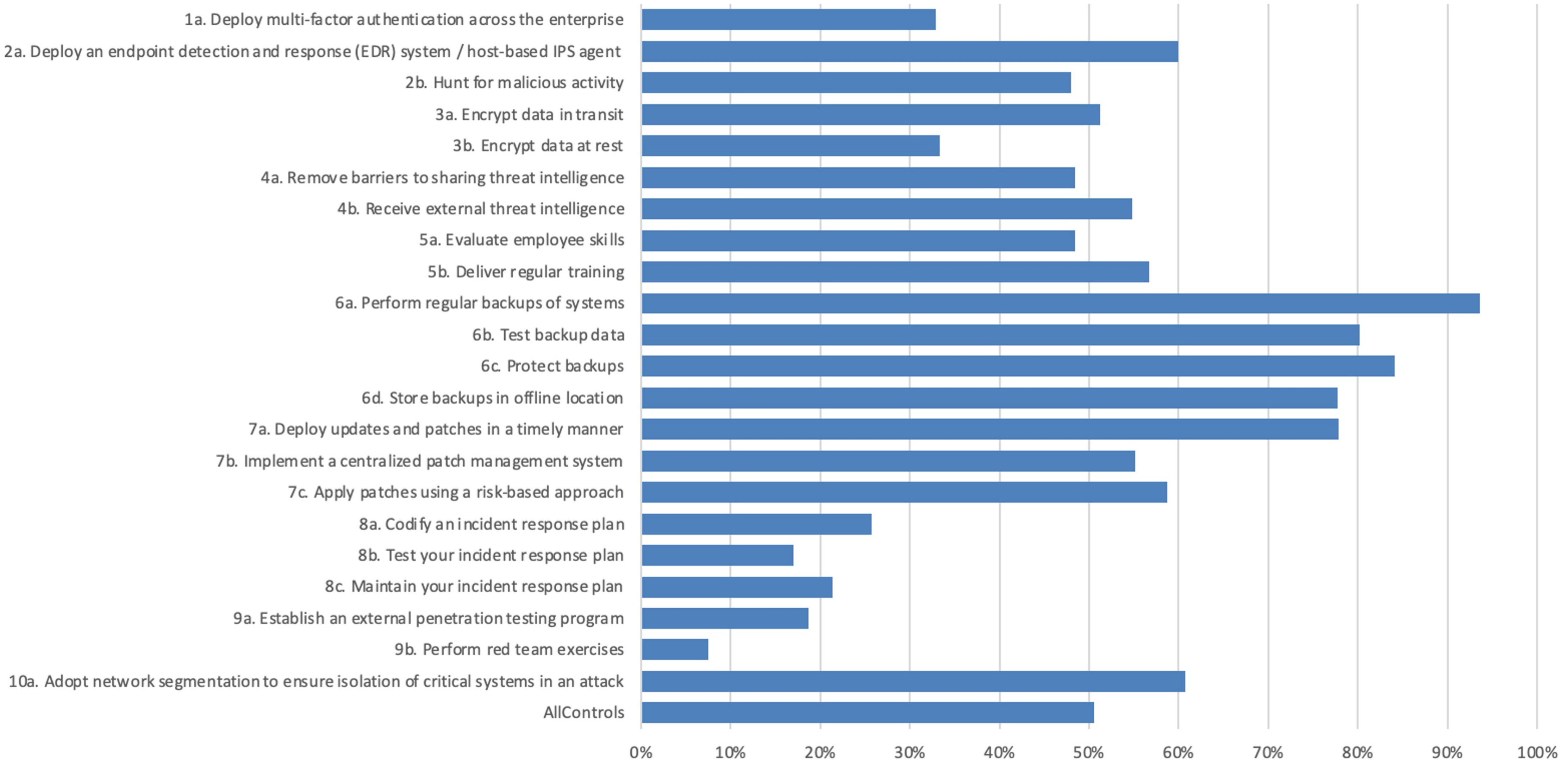
MIT Ransomware Readiness Index (RRI)

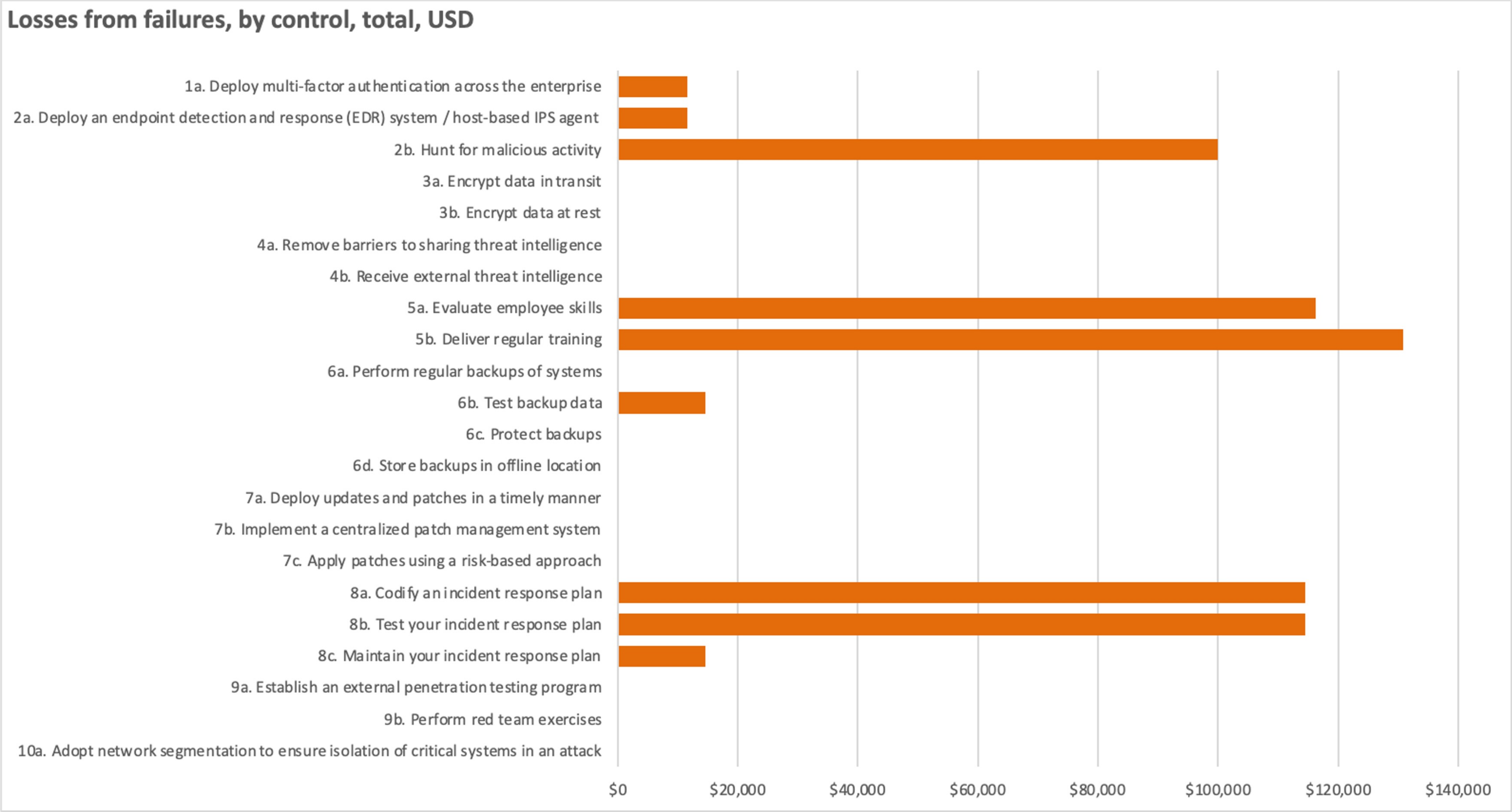


A. Baral, T. Reynolds, L. Susskind, D. Weitzner, Cyber risk modeling for public policy and enterprise decision-making: A case study in municipal cyber risk measurement (Cybersecurity Law and Policy Scholars Conference, Sept 2023)

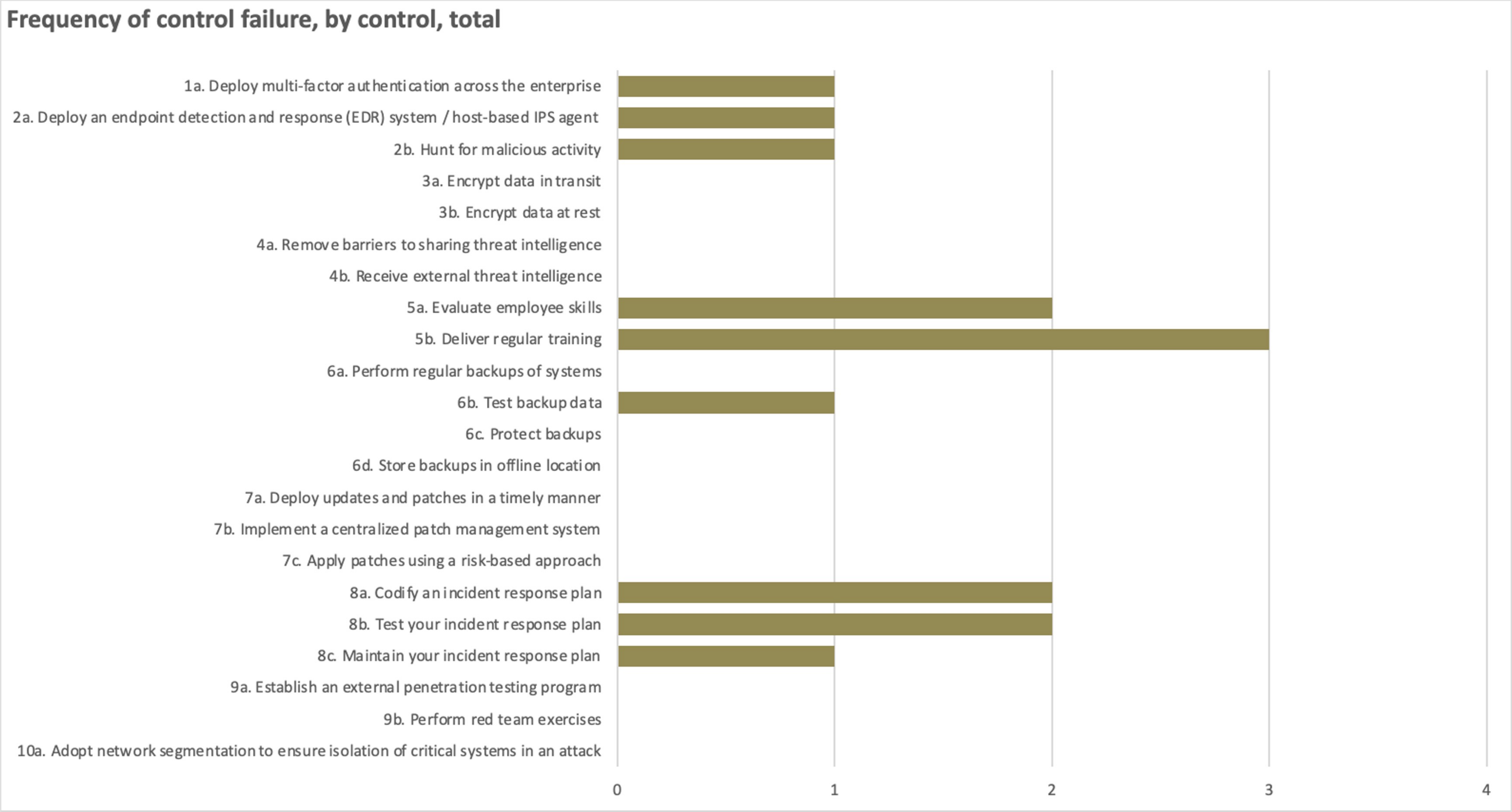
Cyber Risk Measurement for Municipalities

Security control maturity, by control, percentage of municipalities





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
0
1
2
3
4

Evidence-based Policy Guidance

Area	Score	Mat	Adopt	Loss
Incident Resp	28	21	44(a)	20
Check Work	39	14	23(a)	80
Training	50	53	78(a)/35(b)	20
EDR	54	54	69(a)/41(b)	40
MFA	55	33	72(a)/24(b)	60
Encryption	63	42	67(a)	80
Empowerment*	70	52	78(a)	80
Email Security	71		40(b)	
Segment Net	74	61	82(a)	80
Wire Transfer	75		42(b)	
Patch	76	64	84(a)	80
Backup	80	84	96(a)	60

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.

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)

*refers to information sharing and intelligence gathering

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

Calculating an organization's expected risk and optimally efficient investment level

PLG = R

Calculated for the sector (annual P)

Calculated for the sector (average)

Baseline for the sector (average)

Damages limit

P * **L** * **G** = **R**

Probability

Loss

Gap Index (Defense)

Risk

R/P

incident size (forecasted)

Adjusted by firm's actual defenses (1 = no difference)

Firm Specific (annual R)

Firm Specific (incident)

Reasonableness?
If the organization is prepared to cover a loss with a magnitude of R/P, then its actions would be considered reasonable.

Duty of security threshold

Example 1: Muni with **average** security

Calculated for the sector (annual P)		Calculated for the sector (average)		Baseline for the sector (average)				
P	*	**L**	*	**G**	=	**R**	→	R/P
Probability		Loss		Gap Index (Defense)		Risk		Incident size (forecasted)
0.016	*	$157,688	*	1	=	$2,523	→	$157,688
				Adjusted by firm's actual defenses (1 = no difference)		Firm Specific (annual R)		Firm Specific (incident)

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)				
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
				Adjusted by firm's actual defenses (1 = no difference)		Firm Specific (annual R)		Firm Specific (incident)

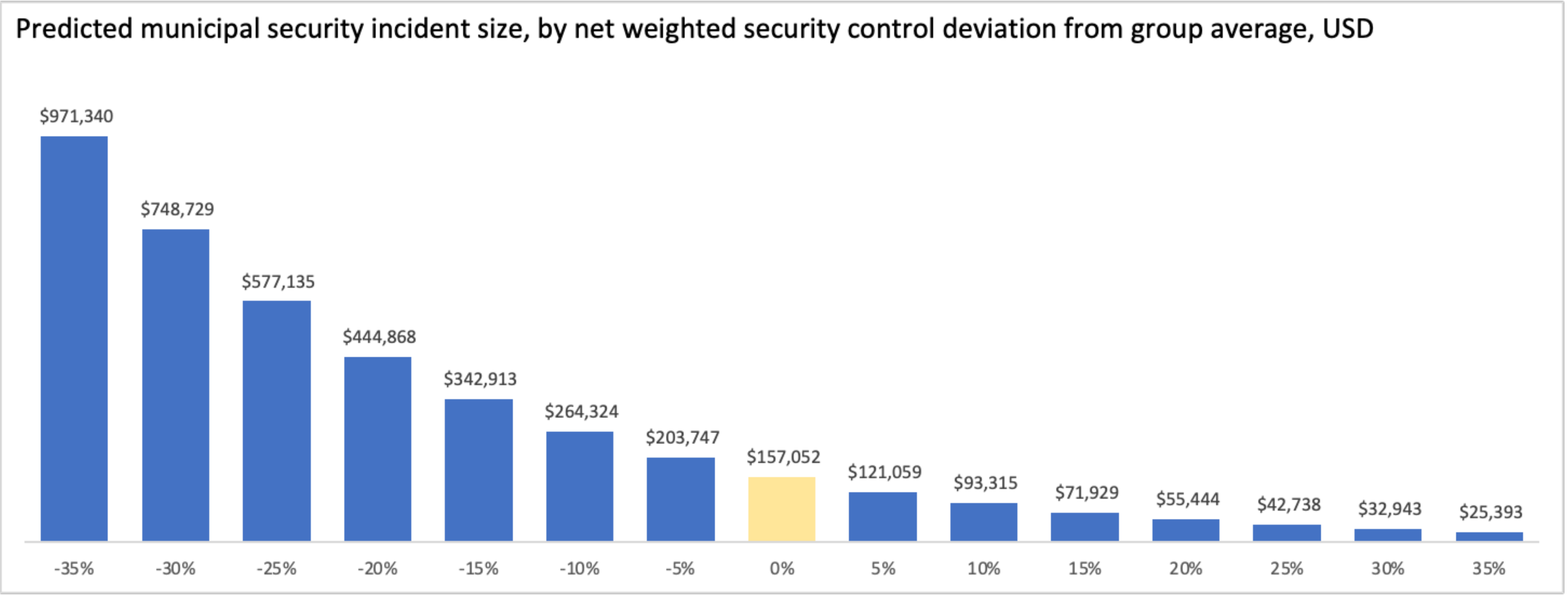
Reasonableness?
If this municipality is prepared to cover a loss of R/P ($748,729), then its actions would be considered reasonable.

Example 3: Muni with **10% higher** security

Calculated for the sector (annual P)		Calculated for the sector (average)		Baseline for the sector (average)				
P	*	**L**	*	**G**	=	**R**	→	R/P
Probability		Loss		Gap Index (Defense)		Risk		Incident size (forecasted)
0.016	*	$157,688	*	0.59	=	$1,493	→	$93,315
				Adjusted by firm's actual defenses (1 = no difference)		Firm Specific (annual R)		Firm Specific (incident)

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

Incident sizes grow quickly with poor security



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.

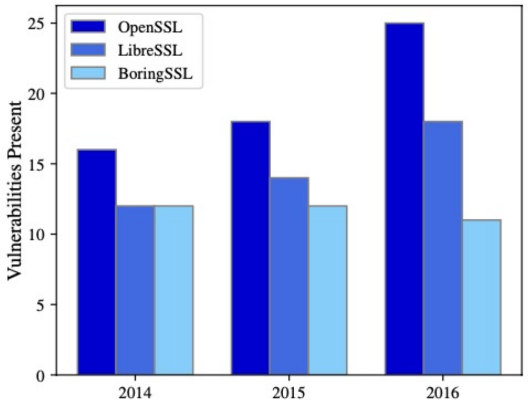


Figure 5: Vulnerability Comparison Post-Heartbleed—Vulnerabilities discovered in OpenSSL after the initial releases of LibreSSL and BoringSSL with a comparison of how many of those vulnerabilities also affected LibreSSL and BoringSSL.

Memory-safe languages produce uniformly more reliable security libraries

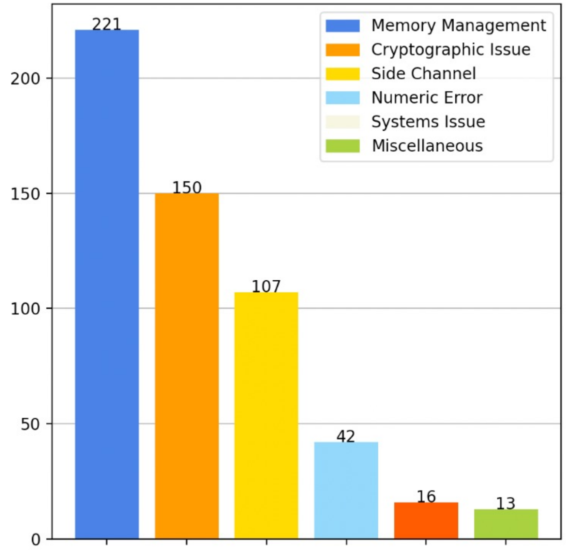


Figure 1: Vulnerability Types — Taxonomy of vulnerabilities by root cause across six main categories within cryptographic software. Table 8 contains descriptions of the taxonomy used and examples of vulnerabilities from each category.

Blessing, Jenny, Michael A. Specter, and Daniel J. Weitzner. "Cryptography in the Wild: An Empirical Analysis of Vulnerabilities in Cryptographic Libraries." Asia CCS 2024 https://arxiv.org/abs/2107.04940

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.

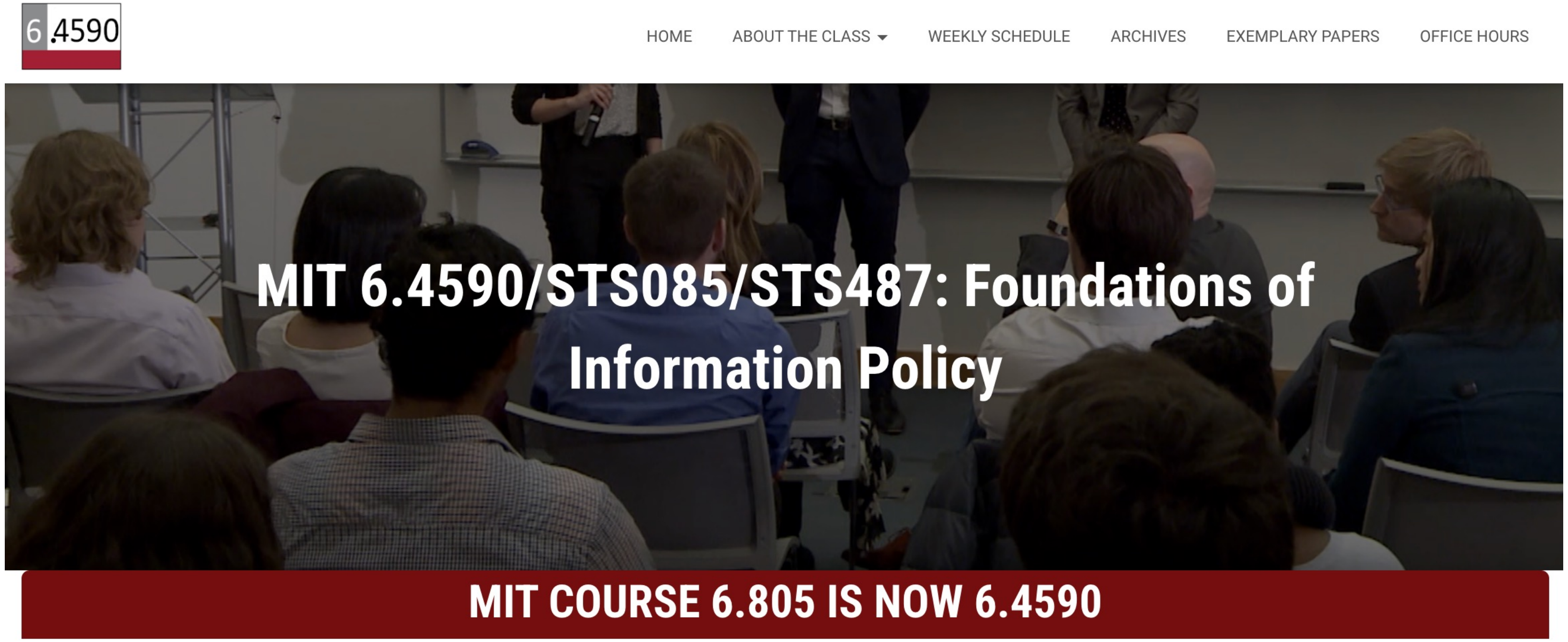
Rationale: In the face of changing threats and vulnerabilities, reasonableness must be dynamically defined. Efficient allocation of cybersecurity resources can only be accomplished based on studying actual losses and control failures as revealed in actual sector-wide and cross-sector data.

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)
- Blessing, Jenny, Michael A. Specter, and Daniel J. Weitzner. "You Really Shouldn't Roll Your Own Crypto: An Empirical Study of Vulnerabilities in Cryptographic Libraries." *arXiv preprint arXiv:2107.04940* (2021)(under review, pre-print available on request)
- Spiewak, Rebecca L., Taylor W. Reynolds, and Daniel J. Weitzner. "MIT Ransomware Readiness Index: A Proposal to Measure Current Preparedness and Progress Over Time." (2021).
- de Castro, L., Lo, A. W., Reynolds, T., Susan, F., Vaikuntanathan, V., Weitzner, D., & Zhang, N. (2020). SCRAM: A Platform for Securely Measuring Cyber Risk. Harvard Data Science Review. https://doi.org/10.1162/99608f92.b4bb506a