Assessing Cybersecurity Risk within the Finance Office

MODERATOR
Steven R. Kreklow
Director, Office of Performance, Strategy & Budget, Milwaukee County Department of Administrative Services

SPEAKER
Douglas W. Hubbard
President, Hubbard Decision Research
How to Measure Anything in Cybersecurity Risk

An Introduction
My Co-Author and I

Richard Seiersen
Currently the General Manager of Cybersecurity and Privacy at GE Health Care. Data driven executive with ~20 years experience spanning subject matters in Cyber Security, Quantitative Risk Management, Predictive Analytics, Big Data and Data Science, Enterprise Integrations and Governance Risk and Compliance (GRC). Led large enterprise teams, provided leadership in multinational organizations and tier one venture capital backed start-ups.

Douglas Hubbard
Uses of Applied Information Economics

AIE was applied initially to IT business cases. But over the last 20 years it has also been applied to other decision analysis problems in all areas of Business Cases, Performance Metrics, Risk Analysis, and Portfolio Prioritization.

<table>
<thead>
<tr>
<th>IT</th>
<th>Business</th>
<th>Government &amp; Non Profit</th>
<th>Military</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Prioritizing IT portfolios</td>
<td>• Movie / film project selection</td>
<td>• Environmental policy</td>
<td>• Forecasting battlefield fuel consumption</td>
</tr>
<tr>
<td>• Risk of software development</td>
<td>• New product development</td>
<td>• Sustainable agriculture</td>
<td>• Effectiveness of combat training to reduce roadside bomb / IED casualties</td>
</tr>
<tr>
<td>• Value of better information</td>
<td>• Pharmaceuticals</td>
<td>• Procurement methods</td>
<td>• R&amp;D portfolios</td>
</tr>
<tr>
<td>• Value of better security</td>
<td>• Medical devices</td>
<td>• Grants management</td>
<td>• Risk of obsolescence and optimal technology upgrades</td>
</tr>
<tr>
<td>• Risk of obsolescence and optimal technology upgrades</td>
<td>• Publishing</td>
<td></td>
<td>• Risk of mine flooding</td>
</tr>
<tr>
<td>• Value of infrastructure</td>
<td>• Real estate</td>
<td></td>
<td>• Risks of major engineering projects</td>
</tr>
<tr>
<td>• Performance metrics for the business value of applications</td>
<td></td>
<td></td>
<td>• Risk of mine flooding</td>
</tr>
</tbody>
</table>
What I Write

• Four books, over 100,000 copies sold in 8 languages.
  – How to Measure Anything: Finding the Value of Intangibles in Business
  – The Failure of Risk Management: Why It’s Broken and How to Fix It
  – Pulse: The New Science of Harnessing Internet Buzz to Track Threats and Opportunities
  – How to Measure Anything in Cybersecurity Risk

• First two are required reading in Society of Actuaries Exam Prep and used in courses at 20+ universities.
Question: What is your single biggest risk in cybersecurity?

Answer: How you measure cybersecurity risk.

(This also applies to risk in general.)
Topics for Today

• Current Methods & Opportunities for Improvement
• How to use expert judgment with “calibration”
• Incremental improvements with decomposition and empirical data
• Summarizing what you can do now, and what you can aspire to do
Topics for Today

• Current Methods & Opportunities for Improvement
  • How to use expert judgment with “calibration”
  • Incremental improvements with decomposition and empirical data
  • Summarizing what you can do now, and what you can aspire to do
Current Solution

Here are some risks plotted on a “typical heat map”.

Suppose mitigation costs were:

- Risk 1: $725K - High
- Risk 2: $95K - Low
- Risk 3: $2.5M - Critical
- Risk 4: $375K - Moderate

What mitigations should be funded and what is the priority among those?
Can Analysis Or Expertise Be A “Placebo”?

“The first principle is that you must not fool yourself, and you are the easiest person to fool.” — Richard P. Feynman

Examples from research:
- Collecting on horse races to predict outcomes (Tsai, Klayman, Hastie)
- Interaction with others to improve project estimates (Heath, Gonzalez)
- Collecting more data about investments to improve returns (Andreassen)

In short, we should assume increased confidence from analysis is a “placebo”. Real benefits have to be measured.
Summarizing Research on Risk Matrices


  – “The burden of proof is squarely on the shoulders of those who would recommend the use of such methods to prove that these obvious inconsistencies do not impair decision making, much less improve it, as is often claimed.’

• Tony Cox “What’s wrong with Risk Matrices” investigates various mathematical consequences of ordinal scales on a matrix.

  – “…they can be “worse than useless,” leading to worse-than-random decisions.”
Is Risk Analysis Actually Supporting Decisions?

- If risks and mitigation strategies were quantified in a meaningful way, decisions could be supported.
- In order to compute an ROI on mitigation decisions, we need to quantify likelihood, monetary impact, cost, and effectiveness.

<table>
<thead>
<tr>
<th>Action</th>
<th>Expected Loss/Yr</th>
<th>Cost of Control</th>
<th>Control Effectiveness</th>
<th>Return on Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB Access</td>
<td>$24.7M</td>
<td>$800K</td>
<td>95%</td>
<td>2,832%</td>
</tr>
<tr>
<td>Mitigate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Access</td>
<td>$2.5M</td>
<td>$300K</td>
<td>99%</td>
<td>727%</td>
</tr>
<tr>
<td>Mitigate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data in Transit</td>
<td>$2.3M</td>
<td>$600K</td>
<td>95%</td>
<td>267%</td>
</tr>
<tr>
<td>Mitigate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network Access Control</td>
<td>$2.3M</td>
<td>$400K</td>
<td>30%</td>
<td>74%</td>
</tr>
<tr>
<td>Mitigate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>File Access</td>
<td>$969K</td>
<td>$600K</td>
<td>90%</td>
<td>45%</td>
</tr>
<tr>
<td>Monitor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web Vulnerabilities</td>
<td>$409K</td>
<td>$800K</td>
<td>95%</td>
<td>-51%</td>
</tr>
<tr>
<td>Track</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System Configuration</td>
<td>$113K</td>
<td>$500K</td>
<td>100%</td>
<td>-77%</td>
</tr>
<tr>
<td>Track</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
How to Build a Method That Works

Start with components that work – that is, the improvement has been measured in controlled tests.

– We can still rely on expert judgement, but it must be calibrated to account for various known errors in expert judgement.

– We have to do the math right and do it with real probabilities including how to update probabilities with empirical data.

If you can’t answer “What is the probability of losing more than X in the next 12 months due to event Y?” then you aren’t doing risk analysis.
One Thing We Know Works Well: Doing the Math

To many experts, when assessing probabilities many events “. . .are perceived as so unique that past history does not seem relevant to the evaluation of their likelihood.” Tversky, Kahneman, *Cognitive Psychology* (1973)

Yet, Historical models routinely outperform experts in a variety of fields (even considering “Black Swans”)

Paul Meehl assessed 150 studies comparing experts to statistical models in many fields (sports, prognosis of liver disease, etc.).

“There is no controversy in social science which shows such a large body of qualitatively diverse studies coming out so uniformly in the same direction as this one.”

Philip Tetlock tracked a total of over 82,000 forecasts from 284 political experts in a 20 year study covering elections, policy effects, wars, the economy and more.

“It is impossible to find any domain in which humans clearly outperformed crude extrapolation algorithms, less still sophisticated statistical ones.”
Watch for Unintentionally Ironic Statements

Have you (or said) heard any of these?

“We don’t have sufficient data, therefore we have to rely on experience.”

All methods of estimating, including expert intuition, have a measurable performance.

“We don’t have sufficient data for a probabilistic analysis, therefore we have to rely on the best estimates of experts.”

Expert judgment must also be based on observation – but usually with faulty recall and inferences.

“Each situation is so unique and complex, that scientific analysis of data is not useful. We are better off relying on our decades of experience.”

Scientific method is based on observation. Therefore, if science is impossible, experience is impossible.
Intuition About Sample Information Is Often Wrong

• Cybersecurity experts are not immune to widely held misconceptions about probabilities and statistics – especially if they vaguely remember some college stats.

• These misconceptions lead many experts to believe they lack data for assessing uncertainties or they need some ideal amount before anything can be inferred.

“Our thesis is that people have strong intuitions about random sampling…these intuitions are wrong in fundamental respects…[and] are shared by naive subjects and by trained scientists”

Amos Tversky and Daniel Kahneman, *Psychological Bulletin*, 1971
The Main Obstacle to Quantitative Methods

Another finding in the same survey: Strong opinions against “quant” are associated with poor stats understanding.

“It’s not what you don’t know that will hurt you, it’s what you know that ain’t so.”

Mark Twain
What Measuring Risk Looks Like

What if we could measure risk more like an actuary – “The probability of losing more than $10 million due to security incidents in 2016 is 16%”

What if we could prioritize security investments based on a “Return on Mitigation”?

<table>
<thead>
<tr>
<th>Action</th>
<th>Expected Loss/Yr</th>
<th>Cost of Control</th>
<th>Control Effectiveness</th>
<th>Return on Control</th>
<th>Expected Loss/Yr</th>
<th>Cost of Control</th>
<th>Control Effectiveness</th>
<th>Return on Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB Access</td>
<td>$24.7M</td>
<td>$800K</td>
<td>95%</td>
<td>2,832%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Access</td>
<td>$2.5M</td>
<td>$300K</td>
<td>99%</td>
<td>727%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data in Transit</td>
<td>$2.3M</td>
<td>$600K</td>
<td>95%</td>
<td>267%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network Access Control</td>
<td>$2.3M</td>
<td>$400K</td>
<td>30%</td>
<td>74%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>File Access</td>
<td>$969K</td>
<td>$600K</td>
<td>90%</td>
<td>45%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web Vulnerabilities</td>
<td>$409K</td>
<td>$800K</td>
<td>95%</td>
<td>-51%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System Configuration</td>
<td>$113K</td>
<td>$500K</td>
<td>100%</td>
<td>-77%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This means there is about a 40% chance of losing more than $10M in a year and about a 10% chance of losing more than $200M.
A Simple “One-For-One Substitution”

Each of these examples can be found on www.howtomeasureanything.com/cybersecurity

<table>
<thead>
<tr>
<th>Event</th>
<th>Event Probability (per Year)</th>
<th>Impact (90% Confidence Interval)</th>
<th>Random Result (zero when the event did not occur)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
</tr>
<tr>
<td>AA</td>
<td>.1</td>
<td>$50,000</td>
<td>$500,000</td>
</tr>
<tr>
<td>AB</td>
<td>.05</td>
<td>$100,000</td>
<td>$10,000,000</td>
</tr>
<tr>
<td>AC</td>
<td>.01</td>
<td>$200,000</td>
<td>$25,000,000</td>
</tr>
<tr>
<td>AD</td>
<td>.03</td>
<td>$100,000</td>
<td>$15,000,000</td>
</tr>
<tr>
<td>AE</td>
<td>.05</td>
<td>$250,000</td>
<td>$30,000,000</td>
</tr>
<tr>
<td>AF</td>
<td>.1</td>
<td>$200,000</td>
<td>$2,000,000</td>
</tr>
<tr>
<td>AG</td>
<td>.07</td>
<td>$1,000,000</td>
<td>$10,000,000</td>
</tr>
<tr>
<td>AH</td>
<td>.02</td>
<td>$100,000</td>
<td>$15,000,000</td>
</tr>
</tbody>
</table>

| ZM    | .05                         | $250,000    | $30,000,000  | 0                                                |
| ZN    | .01                         | $1,500,000  | $40,000,000  | 0                                                |

**Total:** $23,345,193

Each “Dot” on a risk matrix can be better represented as a row on a table like this.

The output can then be

Spreadsheet Example
Topics for Today

• Current Methods & Opportunities for Improvement

• How to use expert judgment with “calibration”

• Incremental improvements with decomposition and empirical data

• Summarizing what you can do now, and what you can aspire to do
Improving Expert Judgement

• Calibration of experts for overconfidence and inconsistency is a start.
• Decomposition tends to further improve expert estimates.
• We can leverage these facts for making improved models even without other recorded, empirical data (adding that comes next).
Undoing Misconceptions About Measurement

- **Measurement**: A quantitatively expressed reduction in uncertainty based on observation.
- You can quantify your current uncertainty (we argue for the “Bayesian” school of thought as the most practical for decisions).
- Even marginal reductions in uncertainty can be extremely valuable.

It’s not a point value.

![Probability Distribution Before Measurement](image1)

![Probability Distribution After Measurement](image2)

**Quantity of Interest**
Quantifying Your Current Uncertainty

• Decades of studies show that most managers are statistically “overconfident” when assessing their own uncertainty.
• Studies also show that measuring your own uncertainty about a quantity is a general skill that can be taught with a measurable improvement.
• Training can “calibrate” people so that of all the times they say they are 90% confident, they will be right 90% of the time.
• HDR has calibrated over 1,000 people in the last 20 years – 85% of participants reach calibration within a half-day of training.

“Overconfident professionals sincerely believe they have expertise, act as experts and look like experts. You will have to struggle to remind yourself that they may be in the grip of an illusion.”

Daniel Kahneman, Psychologist, Economics Nobel
Overconfidence in Ranges

- The same training methods apply to the assessment of uncertain ranges for quantities like the duration of a future outage, the records compromised in a future breach, etc.

![Graph showing overconfidence in ranges](image)

- **Binary Events**: (It happens or not, like a chance of data breach)
  - Ideal Calibration
  - Average of Un-calibrated

- **90% Confidence Interval**: (For continuous values, like impact)
  - Initial
  - Realistic
Calibrating Expert Consistency

- We have gathered over 30,000 individual estimates of probabilities of cyber events from Subject Matter Experts (SMEs).
- Unknown to the SMEs, these estimates included over 2,000 duplicate scenarios pairs.

Comparison of 1\textsuperscript{st} to 2\textsuperscript{nd} Estimates of Cyber risk judgements by same SME

21% of variation in expert responses are explained by \textit{inconsistency}. (79% are explained by the actual information they were given)
A method of improving expert estimates of various quantities was developed in the 1950’s by Egon Brunswik. He called it the “Lens Method.” It has been applied to several types of problems, including expert systems, with consistently beneficial results.
Calibrating Risk Tolerance

- Studies have shown risk aversion changes due to what should be irrelevant external factors including:

<table>
<thead>
<tr>
<th>Factor</th>
<th>Risk Aversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Being around smiling people</td>
<td></td>
</tr>
<tr>
<td>Recalling an event causing fear</td>
<td>↑</td>
</tr>
<tr>
<td>Recalling an event causing anger</td>
<td>↓</td>
</tr>
<tr>
<td>A recent win in an unrelated decision</td>
<td>↓</td>
</tr>
<tr>
<td>A recent loss in an unrelated decision</td>
<td>↑</td>
</tr>
</tbody>
</table>
Loss Exceedance Curves: Before and After

How do we show the risk exposure after applying available mitigations?

- Inherent Risk
- Risk Tolerance
- Residual Risk

Chance of Loss or Greater

Given Loss or Greater (Millions)
Topics for Today

• Current Methods & Opportunities for Improvement
• How to use expert judgment with “calibration”
• Incremental improvements with decomposition and empirical data
• Summarizing what you can do now, and what you can aspire to do
What Published Research Says
(See sources slide for details)

- Psychologists showed that simple decomposition greatly reduces estimation error for estimating the most uncertain variables.
- In the oil industry there is a correlation between the use of quantitative risk analysis methods and financial performance.
- Data at NASA from over 100 space missions showed that Monte Carlo simulations and historical data beat softer methods for estimating cost and schedule risks.
Informative Decompositions

Informative decompositions use what you know or data you can get to improve estimates in models.

**Informative Decompositions:**

- **Systems:** you have fairly detailed knowledge of your applications, what data they have and the hardware it runs on. Some of the parameters of these systems would change your estimate of a risk.

- **Types of Impacts:** You separate confidentiality, integrity and availability events. You have an idea of business volumes like sales and other processes. If a breach or outage occurred, you can describe something about the consequences.

- **Staff:** You have knowledge of the number of employees, device loss rates, and some knowledge of what data they may have.

- **Vendors & Customers:** You know who the parties you interact with and you have some knowledge about them.

- **Insurance:** Any cyber-insurance will have detailed language regarding limitations, exclusions, etc.
Measurement Challenge: Reputation Damage

- One of the perceived most difficult measurements in cybersecurity is damage to reputation.
- Trick: *There is no such thing as a “secret” damage to reputation!*
- How about comparing stock prices after incidents? (That’s all public!)
- So what is the *REAL* damage?
  - Legal liabilities,
  - Customer outreach
  - “Penance” projects (security overkill)
- The upshot, damage to reputation actually has available information and easily observable measured costs incurred to *avoid* the bigger damages!
Statistics Needs Less Data Than You Think

- You have relatively few examples of major, reported breaches in each industry.
- There is a statistical method for estimating the frequency of breaches based on small samples. This is the “beta” distribution and it is provided in Excel as “=betadist(proportion, hits, misses)”.
- Spreadsheet for this at www.howtomeasureanything.com/cybersecurity

Out of 98 Retail had 3 breaches from Jan 2014 to June 2015

Annual Breach Frequency per Organization

Spreadsheet Example
What Reduces Data Breach Risk?

- The survey reveals another interesting result.
- Those who said they computed the probability of losses reported fewer breaches than those who did not.
- I would not treat this observation alone as sufficient – but it agrees with other independent evidence.

Participants who said YES to "*We are able to compute the probability of various levels of losses for the organization.*"

Results indicate a **97% chance** that those who answered YES have lower breach frequency.
Data from the HHS “Wall of Shame” indicates that the rate of data breaches (more than 500 confidential records) is now consistently 14% per year per 10,000 employees.

Source: Vivosecurity
Topics for Today

• Current Methods & Opportunities for Improvement
• How to use expert judgment with “calibration”
• Incremental improvements with decomposition and empirical data
• Summarizing what you can do now, and what you can aspire to do
What to Do Next

Things you can do now:
• Stop using ordinal scales and risk matrices for evaluating risk – they only create the illusion of analysis
• Replace risk matrix activities with the One-for-One Substitution model
• Experiment with simple additional decompositions as shown in the Chapter 6 download (both spreadsheets available at www.hubbardresearch.com/cybersecurity)

Things to strive toward (the effort is easily justified for Cybersecurity):
• Get calibrated so you can quantify your uncertainty
• Learn more advanced decompositions including Log Odds and the Lens Method
• Update the initial model with empirical data using slightly more advanced statistical methods

Measure what matters. Make better decisions.
A Major Fallacy In Comparing Methods

• Don’t make the classic “Beat the Bear” fallacy.

Exsupero Ursus

• If you doubt the effectiveness of quantitative methods, remember, all you have to do is outperform the alternative:

• …unaided expertise or soft scoring methods.
Questions?

Contact:
Doug Hubbard
Hubbard Decision Research
dwhubbard@hubbardresearch.com
www.hubbardresearch.com
630 858 2788